

RESULT 33  
 ID 09LNS7 PRELIMINARY; PRT; 651 AA.  
 AC 09LNS7;  
 DT 01-OC7-2000 (TREMblrel. 15, Created)  
 DT 01-OC7-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE T20H2.23 protein.  
 GN T20H2.23.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.  
 NCBI\_TaxID=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Sakano H., Vaysberg M., Lee J.M., Lenz C., Liu S., Pham P.,  
 RA Toriumi M., Yu G., Chin C., Chlou J., Choi E., Chung M., Gonzalez A.,  
 RA Hwang B., Liu A., Altafi H., Brooks S., Buehler E., Chao Q., Conn L.,  
 RA Conway A.B., Hansen N.F., Johnson-Hopson C., Khan S., Kim C., Lam B.,  
 RA Miranda M., Nguyen M., Palm C.J., Shinn P., Southwick A., Davis R.W.,  
 RA Ecker J.R., Federspiel N.A., Theologis A.;  
 RA "The sequence of BAC T20H2 from Arabidopsis thaliana chromosome 1.";  
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AC022472; AAF7914.1;  
 DR EMBL: AC022472; AAF7914.1;  
 SO SEQUENCE 651 AA; 73831 MW; 8AB2D146004F7896 CRC64;

Query Match 7.2%; Score 7; DB 10; Length 651;  
 Best local Similarity 100.0%; Pred. No. 46;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 KKKRERK 34  
 Db 507 KKKRERK 513

RESULT 34  
 ID 09V6S8 PRELIMINARY; PRT; 660 AA.  
 AC 09V6S8;  
 DT 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMblrel. 13, Last annotation update)  
 DE CG13337 protein.  
 GN CG13337.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Georger R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Brandon R.C., Rogers Y.H.C., Blake J.R.G., Chapple M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Abmayyan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.F., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson M.K., Nixon K., Nusskern D.R., Pauley J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Pollard J., Reese M.G.,  
 RA Reinert K., Remington K., Sanders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton K., Strong R., Sun E.,  
 RA Swikdas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AEO03817; AAF58344.1;  
 DR EMBL: AEO03817; AAF58344.1;  
 DR FLYBASE; FBgn0033663; CG13337.  
 SO SEQUENCE 660 AA; 79758 MW; E10B28EFBE127FAD CRC64;

Query Match 7.2%; Score 7; DB 5; Length 680;  
 Best local Similarity 100.0%; Pred. No. 48;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 26 ERKKRE 32  
 Db 461 ERKKRE 467

RESULT 35  
 ID 015816 PRELIMINARY; PRT; 699 AA.  
 AC 015816;  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
 DE MYD2 protein.  
 GN MYD2.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.  
 NCBI\_TaxID=44689;  
 RN  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX3;  
 RX MEDLINE=98283920; PubMed=9620859;  
 RA Otsuka H., Van Haastert P.J.M.;  
 RA "A novel Myb homologue initiates Dictyostelium development by  
 induction of adenylate cyclase expression.";  
 RT Genes Dev. 12:1738-1748(1998).  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 3 MYB-LIKE DOMAINS.  
 DR EMBL: AJ002383; CAA05357.1;  
 DR HSPB; P06876; 1MR8;  
 DR InterPro; IPR001005; MYB\_DNA\_binding.  
 DR Pfam; PF00249; MYB\_DNA\_binding; 3.  
 DR SMART; SM00395; SANT; 3.  
 DR PROSITE; PS50090; MYB\_3; 3.  
 DR DNA-binding; Nuclear protein.  
 SO SEQUENCE 699 AA; 80381 MW; EA989DC77953F265 CRC64;

Query Match 7.2%; Score 7; DB 5; Length 699;  
 Best local Similarity 100.0%; Pred. No. 49;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 28 KKKRERK 34  
 Db 409 KKKRERK 415

```

RESULT 36
Q9R2F3      PRELIMINARY:      PRT:      703 AA.
ID 09R2F3;
AC 09R2F3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 10, Last annotation update)
DE Nduvulation protein-related protein.
GN DUC0037.
OS Deinococcus radiodurans.
OC Plasmid Cpl.
OC Bacteria: Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Uitterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1";
RL Science 286:1571-1577(1999).
DR EMBL; AE001827; AAF12674.1; -
DR TIGR; DRC0037; -
DR InterPro; IPR003696; Carbintransf.
DR Pfam; PF02543; CmcH_NodJ; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 703 AA; 77447 MW; 6ED2493D0ECD46 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 16; Length 703;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 12 GDSSTLS 18
DB 165 GDSSTLS 171

RESULT 37
Q9H1V1      PRELIMINARY:      PRT:      772 AA.
ID 09H1V1;
AC 09H1V1;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
DE DJ127C7.1 (K1AA0250 protein).
GN DJ127C7.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DE EMBL; AL137800; CAC19685.1; -
SQ SEQUENCE 772 AA; 85964 MW; FED3650A34447037 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 4; Length 772;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 SOTELRK 24
DB 282 SOTELRK 288

```

```

RESULT 38
O81868      PRELIMINARY:      PRT:      777 AA.
ID O81868;
AC O81868;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypothetical 87.4 kDa protein.
GN T16H5.230 OR AT4G19870.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicot; Rosidae;
OC Eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA De Haan M., Maarse A.C., Grivell L.A., Bancroft I., Mewes H.W.,
RA Mayer K., Schueller C., Bevan M.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL024486; CAA19704.1; -
DR EMBL; AL161551; CAB78969.1; -
DR InterPro; IPR001810; F-box.
DR InterPro; IPR001798; Kelch.
DR Pfam; PF00646; F-box; 2.
DR Pfam; PF01344; Kelch; 4.
DR PRINTS; PR00501; KELCHREPEAT.
DR SMART; SM00256; FBOX; 2.
DR PROSITE; PSS0181; FBOX; 2.
KW Hypothetical protein.
SQ SEQUENCE 777 AA; 87450 MW; 1ABCC50FD6DED789 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 10; Length 777;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 86 ATRSHLG 92
DB 455 ATRSHLG 461

RESULT 39
Q96N31      PRELIMINARY:      PRT:      799 AA.
ID Q96N31;
AC Q96N31;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE CDNA FLJ31473 f1s, clone NT2NE2001530 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RA Nishimura K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project";
SQ Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

```

RT "Genome sequence of the H1N1 virus"

```

RT Investigating biology "The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Nham M.;
RT "The sequence of C. elegans cosmid F59A6.";
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U41994; AAK31526.1; -.
DR InterPro: IPR001180; CItrom.
DR InterPro: IPR002219; DAG_PE-blind.
DR Pfam: PF00780; CNH; 1.
DR Pfam: PF00169; PH; 1.
DR SMART: SM00109; CL; 1.
DR SMART: SM00036; CNH; 1.
DR SMART: SM00233; PH; 1.
DR PROSITE: PSS0081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE: PSS0003; PH_DOMAIN; 1.
SQ SEQUENCE 1286 AA; 148666 MW; 1402C3A80DC12BE4 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 5; Length 1286;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 ASIGDSE 15
DB 317 ASIGDSE 323

RESULT 44
Q22944 PRELIMINARY; PRT; 1336 AA.
AC Q22944;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE C52B9.8 protein.
GN C52B9.8.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peleoderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology the C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Nelson J.;
RT "The sequence of C. elegans cosmid C52B9.";
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U64598; AAK39219.1; -.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR001650; Helicase_C.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00176; SNF2_N; 1.

PRINTS: PR01632; POYDCALPHAL.
DR SMART: SM00487; DEAD; 1.
DR SMART: SM00490; HELIC; 1.
KW ATP-binding; Helicase.
SQ SEQUENCE 1336 AA; 156127 MW; DBADE8CFAE0593B8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 5; Length 1336;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 QTELRRK 25
DB 134 QTELRRK 140

RESULT 45
Q03291 PRELIMINARY; PRT; 1435 AA.
AC Q03291;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JUN-1996 (TREMBlrel. 01, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
GN SPP41 OR YDR464W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RP SEQUENCE FROM N.A.
RA Dietrich F.S.; Mulligan J., Allen E., Araujo R., Aviles E., Berno A.,
RA Carpenter J., Chen E., Cherry J.M., Chung E., Duncan M.,
RA Hunkle-Smith S., Hyman R., Komp C., Lashkari D., Lew H., Lin D.,
RA Moseedale D., Nakahara K., Namath A., Oefner P., Oh C., Petel F.X.,
RA Roberts D., Schramm S., Schroeder M., Shogren T., Shroff N.,
RA Winant A., Yellon M., Botstein D., Davis R.W.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RP SEQUENCE FROM N.A.
RA Dietrich F.S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Jia Y., Cherry J.M.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U33050; AAB64901.1; -.
DR SCD; S0002872; SPP41.
DR InterPro: IPR003903; UTM.
SQ SEQUENCE 1435 AA; 161596 MW; 7C18CF349FD647E8 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 3; Length 1435;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 KKERKK 30
DB 695 KKERKK 701

Search completed: May 11, 2003, 20:12:09
Job time: 56.5044 secs

```



GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:05:46 ; Search time 2182.93 Seconds  
(without alignments)  
1293.204 Million cell updates/sec

Title: US-09-854-133-586  
Perfect score: 1 EVEVSNDHSLGDSFSLSQT.....ITGCLPMATRSHTGRKCS 97  
Sequence:

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4106486

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: listing first 60 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09854133/runat\_05052003\_174132\_704/app\_query.fasta.1.462  
-DB=genemb1 -OEMT=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human4.0.cdi -LIST=60  
-DOCALLIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=45 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09854133.ecgn.1.1758-etunal\_05052003\_174132\_704 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMOUT=120  
-MARN\_TIMOUT=30 -TREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database:

GenEmb1:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vt:\*  
15: gb\_da:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_cm:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vt:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_inv:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rdt:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	337	6 AX321911	AX321911 Sequence
2	96	99.0	5981	6 AX321910	AX321910 Sequence
3	93	95.9	2239	6 AX321909	AX321909 Sequence
4	60	61.9	161280	9 AC093903	AC093903 Homo sapi
5	37	38.1	575	6 AX341060	AX341060 Sequence
6	37	38.1	596	6 AX351341	AX351341 Sequence
7	37	38.1	1683	9 AB042201	AB042201 Homo sapi
8	37	38.1	1861	9 AB026891	AB026891 Homo sapi
9	37	38.1	1874	9 AF200708	AF200708 Homo sapi
10	37	38.1	2000	9 AB040875	AB040875 Homo sapi
11	37	38.1	2155	9 BC012087	BC012087 Homo sapi
12	37	38.1	2482	9 AF252872	AF252872 Homo sapi
13	37	38.1	3144	9 HSA27882	HSA27882 Homo sapi
14	12	12.4	295	5 TCR292491	TCR292491 Human DNA
15	12	12.4	681	4 CH19F15314	CH19F15314 Human DNA
16	12	12.4	41362	9 AC101290	AC101290 Mus muscu
17	12	12.4	68527	2 AC022364	AC022364 Homo sapi
18	12	12.4	7537	2 AF000960	AF000960 Homo sapi
19	12	12.4	76163	9 HSI18963	HSI18963 Human DNA
20	12	12.4	100167	9 AC090033	AC090033 Felis cat
21	12	12.4	125267	4 AC090033	AC090033 Felis cat
22	12	12.4	125661	2 AC087731	AC087731 Felis cat
23	12	12.4	127475	2 AC105404	AC105404 Mus muscu
24	12	12.4	143786	2 AC015503	AC015503 Homo sapi
25	12	12.4	145722	2 AC004932	AC004932 Homo sapi
26	12	12.4	151183	2 AC011275	AC011275 Homo sapi
27	12	12.4	152959	2 AC011492	AC011492 Homo sapi
28	12	12.4	156795	2 AC022290	AC022290 Homo sapi
29	12	12.4	156795	2 AC022290	AC022290 Homo sapi
30	12	12.4	163980	2 AC120209	AC120209 Felis cat
31	12	12.4	170513	2 AC023603	AC023603 Homo sapi
32	12	12.4	171506	2 AC120322	AC120322 Rattus no
33	12	12.4	172571	2 AC027484	AC027484 Mus muscu
34	12	12.4	173700	2 AC116823	AC116823 Homo sapi
35	12	12.4	173735	2 AC105990	AC105990 Mus muscu
36	12	12.4	175398	2 AC027063	AC027063 Homo sapi
37	12	12.4	175690	9 AC009989	AC009989 Homo sapi
38	12	12.4	177444	9 AC011468	AC011468 Homo sapi
39	12	12.4	178361	9 AC107305	AC107305 Homo sapi
40	12	12.4	179343	2 AC121904	AC121904 Mus muscu
41	12	12.4	180336	2 AC108976	AC108976 Rattus no
42	12	12.4	181445	9 AC026367	AC026367 Homo sapi
43	12	12.4	183011	9 AC118757	AC118757 Homo sapi
44	12	12.4	186298	9 AC093768	AC093768 Homo sapi
45	12	12.4	186920	9 AC121187	AC121187 Rattus no
46	12	12.4	188690	2 AC122690	AC122690 Homo sapi
47	12	12.4	189169	2 AC121771	AC121771 Mus muscu
48	12	12.4	193171	2 AC121771	AC121771 Mus muscu
49	12	12.4	193171	2 AC121771	AC121771 Mus muscu
50	12	12.4	193171	2 AC121771	AC121771 Mus muscu





Alignment Scores:	
Pred. No.:	2.03e-52
Score:	60.0%
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	61.86%
DB:	9
	Gaps:
	0
US-09-854-133-586 (1-97) x AC0093903 (1-161280)	
	Length: 161280
	Matches: 60
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

QY	1	GLUVALGLVALSERARGASPHLSALAISERLEUGLYASPSERGIDPHRLSEUGLINTHR	20
Db	27755	GAGGTGGAGTGGACGAGAGATCATGCACCGCTGGGTGACAGTAGACTGTCTCAACA	27696
QY	21	GLULEAURLYLSGLIURLYLYLSLYLSARGGLIURLYSPHEGLIALAASCYGLY	40
Db	27695	GAAATTAAGCAAAAAAAGAAAGAAAGAAAAAGAGAGAGAGAAATTCAGGCCCAATTGTGGC	27636
QY	41	ILEASPHEILEILEPHERPLIEPHERPLIEULEUPHESERHISHSTRIPLEGIN	60
Db	27635	ATAGATTTTATCATATTCTCGAATTTTGTGGATTTCTTTTGTTCATCATCATGAGATTGAG	27576
RESULT 5	AX341060	Sequence 1307 from Patent WO0196388.	DNA linear PAT 10-JAN-2002
LOCUS	AX341060		
DEFINITION	Sequence 1307 from Patent WO0196388.		
ACCESSION	AX341060		
VERSION	AX341060.1	GI:18137042	
FEATURES			

SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	FEATURES	SOURCE
human.	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	Jiang, Y., Harlocker, S.L. and Secrist, H.	Compositions and methods for the therapy and diagnosis of colon cancer	Patent: WO 0196388-A 1307 20-DIC-2001; CORIXA CORPORATION (US)	1	575
					Location/Qualifiers		

BASE COUNT	151 a	122 c	153 g	141 t	8 others
ORIGIN					
Alignment Scores:					
Pred. No.:	4.38e-30	Length:			
Score:	37.00	Matches: 575			
Percent Similarity:	100.00%	Conservative: 37			
Best Local Similarity:	100.00%	Mismatch: 0			
Query Match:	38.14%	Indels: 0			
DB:	6				

US-09-854-133-586 (1-97) x AX351341 (1-575)

QY 61 GluSerLeuLeuCysProProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 |||||  
 Db 256 GAAAGCCCTGTGTGTCACCAATCTCCAAAGAGGTTACTGACGAGGAATGTTAACGGGA 315  
 |||||

QY 81 GlyCysLeuProTPALaThrArgSerHisLeuGlyArgArgLyCysSer 97  
 |||||  
 Db 316 GGCTGCTTCTCCCTGGCAACAAGAGCCACTGGGAGAGGAAGTGCAGC 366  
 |||||

RESULT 6  
 AX351341/c 596 bp DNA linear PAT 06-FEB-2002  
 LOCUS Sequence 88 from Patent WO0196390.  
 DEFINITION AX351341  
 ACCESSION AX351341  
 VERSION AX351341.1 GI:18616688  
 KEYWORDS  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 1 Jiang, Y., Hepler, W.T., Clapper, J.D., Wang, A. and Secrist, H.  
 Authors Compositions and methods for the therapy and diagnosis of colon  
 title cancer  
 JOURNAL Patent: WO 0196390-A 88 20-DEC-2001;  
 CORIXA CORPORATION (US)  
 FEATURES Location/Qualifiers  
 source 1..596  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 BASE COUNT 160 a 151 c 126 g 158 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 4,52e-30 Length: 596  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 38.14% Indels: 0  
 DB: Gaps: 6

US-09-854-133-586 (1-97) x AX351341 (1-596)

QY 61 GluSerLeuLeuCysProProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 |||||  
 Db 367 GAAAGCCCTGTGTGTCACCAATCTCCAAAGAGGTTACTGACGAGGAATGTTAACGGGA 308  
 |||||

QY 81 GlyCysLeuProTPALaThrArgSerHisLeuGlyArgArgLyCysSer 97  
 |||||  
 Db 307 GGCTGCTTCTCCCTGGCAACAAGAGCCACTGGGAGAGGAAGTGCAGC 257  
 |||||

RESULT 7  
 AB042201 1683 bp DNA linear PRI 10-FEB-2001  
 LOCUS Homo sapiens xCT gene for cystine/glutamate transporter, exon 1,  
 DEFINITION partial cds.  
 ACCESSION AB042201.1 GI:7670267  
 VERSION AB042201.1  
 KEYWORDS cystine/glutamate transporter.  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 1 (sites)  
 Authors Sato, H., Yamba, M., Kuriyama-Matsumura, K., Okuno, S. and Bannai, S.  
 title Molecular cloning and expression of human xCT, the light chain of  
 amino acid transport system xc-  
 JOURNAL Antioxidants and Redox Signaling 2, 665-671 (2000)  
 REFERENCE 2 (bases 1 to 1683)  
 Authors Sato, H. and Bannai, S.  
 title Direct Submission

JOURNAL Submitted (24-APR-2000) Hideyo Sato, University of Tsukuba,  
 Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba,  
 Ibaraki 305-8575, Japan (E-mail:hideyo-s@nd.tsuakuba.ac.jp,  
 Tel:81-298-53-3282, Fax:81-298-53-3039)

FEATURES  
 source 1..1683  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 gene 1001..1683  
 /gene="xCT"  
 1001..1555  
 /gene="xCT"  
 /product="cystine/glutamate transporter"  
 /number=1  
 1279..>1555  
 /gene="xCT"  
 /codon\_start=1  
 /product="cystine/glutamate transporter"  
 /protein\_id="BA09499.1"  
 /db\_xref="GI:7670268"  
 /translation="MYRKPVSTISKGYLQGNVNGRLPSLGNKEPPGGEYQAKRKV  
 TLLRGSIITGITGAGIFISPKGYLQNTGSYQMSLITVQGVLSLF"  
 1556..>1683  
 /gene="xCT"  
 /number=1  
 /db\_xref="taxon:9606"

BASE COUNT 495 a 287 c 370 g 531 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.11e-29 Length: 1683  
 Score: 37.00 Matches: 37  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 38.14% Indels: 0  
 DB: Gaps: 9

US-09-854-133-586 (1-97) x AB042201 (1-1683)

QY 61 GluSerLeuLeuCysProProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 |||||  
 Db 1286 GAAAGCCCTGTGTGTCACCAATCTCCAAAGAGGTTACTGACGAGGAATGTTAACGGGA 1345  
 |||||

QY 81 GlyCysLeuProTPALaThrArgSerHisLeuGlyArgArgLyCysSer 97  
 |||||  
 Db 1346 GGCTGCTTCTCCCTGGCAACAAGAGCCACTGGGAGAGGAAGTGCAGC 1396  
 |||||

RESULT 8  
 AB026891 1861 bp mRNA linear PRI 10-FEB-2001  
 LOCUS Homo sapiens mRNA for cystine/glutamate transporter, complete cds.  
 DEFINITION AB026891  
 ACCESSION AB026891.1 GI:5668544  
 VERSION AB026891.1  
 KEYWORDS cystine/glutamate transporter.  
 SOURCE Homo sapiens fibroblast cDNA to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
 1 (sites)  
 Authors Sato, H., Yamba, M., Kuriyama-Matsumura, K., Okuno, S. and Bannai, S.  
 title Molecular cloning and expression of human xCT, the light chain of  
 amino acid transport system xc-  
 JOURNAL Antioxidants and Redox Signaling 2, 665-671 (2000)  
 REFERENCE 2 (bases 1 to 1861)  
 Authors Sato, H. and Bannai, S.  
 title Direct Submission  
 JOURNAL Submitted (30-APR-1999) Hideyo Sato, University of Tsukuba,  
 Institute of Basic Medical Sciences; Tennodai 1-1-1, Tsukuba,  
 Ibaraki 305-8575, Japan (E-mail:hideyo-s@nd.tsuakuba.ac.jp,  
 Tel:81-298-53-3282, Fax:81-298-53-3039)

FEATURES  
 source 1..1861  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"

gene  
1. .1861  
/gene="hct"  
236. .1741  
/gene="hct"  
/codon\_start=1  
/product="cystine/glutamate transporter"  
/protein\_id="BA82628.1"  
/db\_xref="GI:5668545"  
/translation="MYRKPVSSTISKGYLQGNVNGRLPSLGNKEPPQGRVQLKRY  
TLRGVSIITIGTIGACIFIPKGVLONTSVGSMITITVCGVLSLFGALSYAELGT  
TIKSGGHTYITIEVGPPLAFVWVWELLIRPATAVISAAGRYILPEFIOCI  
PELAIKLITAVGITVWVWVNSVMSARIQIFLFCCKLTAIILITVPGMOLIKOT  
ONKDAFSGRDSSITRLPLAFYGMVAVAGVYLNFTVEVENPEKTIPLATISMAI  
VTIGVLTNVAFFTTINAEELLSNVAVTFESRLGNFSLAVPIFVALSCFSNMG  
VFASRLFYVASRSGHLPETLSMIVHKKRPPLAVIYVLPIMLIFSGDLSLNL  
SFARWLFGVAVAGLILRYKCPDMRPKVPPLFPLFSPFCMLFVALSLYSDPST  
GIGFVITLGPVAYYLFITMDKPRFMRSEKTRITLQIILEVPEEDKL"

BASE COUNT 485 a 402 c 421 g 553 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.22e-29 Length: 1861  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x AB026891 (1-1861)

Qy 61 Glusertleucysproproserprolysgluvalthrcysargglumethrly 80  
Db 243 GAAAGCCTGTGTGTCACCAATCCAAAGAGAGTTACTCAGGAAATGTTAAACGGA 302

Qy 81 Glycylseuprotropalathrarserhisleuglyargarglyscys 97  
Db 303 GGCTGCTTCCCTGGCAACAGAGCCACTGGCAGAGAGAGAGTGCAGC 353

RESULT 9  
AF200708 1874 bp mRNA linear PRI 01-DEC-2000  
LOCUS Homo sapiens calcium channel blocker resistance protein CCBRL mRNA,  
complete cds.  
ACCESSION AF200708  
VERSION AF200708.1 GI:11493651  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens.  
REFERENCE 1 (bases 1 to 1874)  
AUTHORS Conklin,D.S. and Beach,D.H.  
TITLE CCBRL, novel CD98 light chain implicated in redox control and  
calcium signaling  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1874)  
AUTHORS Conklin,D.S. and Beach,D.H.  
TITLE Direct Submission  
JOURNAL Submitted (01-NOV-1999) Cold Spring Harbor Labs, 1 Bungtown Rd,  
Cold Spring Harbor, NY 11724, USA  
FEATURES  
source location/Qualifiers  
1. .1874  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_line="fibrosarcoma HT1080"  
246. .1751  
/note="X-c transporter; CD98 light chain"  
/codon\_start=1  
/product="calcium channel blocker resistance protein  
CCBRL"  
/protein\_id="AA035592.1"  
/db\_xref="GI:11493652"

/translation="MYRKPVSSTISKGYLQGNVNGRLPSLGNKEPPQGRVQLKRY  
TLRGVSIITIGTIGACIFIPKGVLONTSVGSMITITVCGVLSLFGALSYAELGT  
TIKSGGHTYITIEVGPPLAFVWVWELLIRPATAVISAAGRYILPEFIOCI  
PELAIKLITAVGITVWVWVNSVMSARIQIFLFCCKLTAIILITVPGMOLIKOT  
ONKDAFSGRDSSITRLPLAFYGMVAVAGVYLNFTVEVENPEKTIPLATISMAI  
VTIGVLTNVAFFTTINAEELLSNVAVTFESRLGNFSLAVPIFVALSCFSNMG  
VFASRLFYVASRSGHLPETLSMIVHKKRPPLAVIYVLPIMLIFSGDLSLNL  
SFARWLFGVAVAGLILRYKCPDMRPKVPPLFPLFSPFCMLFVALSLYSDPST  
GIGFVITLGPVAYYLFITMDKPRFMRSEKTRITLQIILEVPEEDKL"

BASE COUNT 490 a 407 c 422 g 555 t

ORIGIN

Alignment Scores:  
Pred. No.: 1.22e-29 Length: 1874  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x AF200708 (1-1874)

Qy 61 Glusertleucysproproserprolysgluvalthrcysargglumethrly 80  
Db 253 GAAAGCCTGTGTGTCACCAATCCAAAGAGAGTTACTCAGGAAATGTTAAACGGA 312

Qy 81 Glycylseuprotropalathrarserhisleuglyargarglyscys 97  
Db 313 GGCTGCTTCCCTGGCAACAGAGCCACTGGCAGAGAGAGAGTGCAGC 363

RESULT 10  
AB040875 2000 bp mRNA linear PRI 03-APR-2001  
LOCUS Homo sapiens hxcT mRNA for cystine/glutamate exchanger, complete  
cds.  
ACCESSION AB040875  
VERSION AB040875.1 GI:13516845  
KEYWORDS  
SOURCE Homo sapiens adult cDNA to mRNA, clone\_1lb:brain cDNA library  
clone:hxcT.  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 2000)  
AUTHORS Kim,J.Y., Chai, Rongduan,A., Cha,S.H., Segawa,H., Matsuo,H.,  
Kim,D.K., Endou,H. and Kanai,Y.  
TITLE Human cystine/glutamate exchanger: cDNA cloning and upregulation by  
oxidative stress in glioma cells  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2000)  
AUTHORS Kanai,Y.  
TITLE Direct Submission  
JOURNAL Submitted (01-APR-2000) Yoshikatsu Kanai, Kyorin University School  
of Medicine, Department of Pharmacology and Toxicology; 6-20-2  
Shinkawa, Mitaka, Tokyo 181-8611, Japan  
(E-mail:ykanai@kyorin-u.ac.jp, Tel:+81-422-47-5511(ex.3453),  
Fax:+81-422-79-1321)  
FEATURES  
source location/Qualifiers  
1. .2000  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="hxcT"  
/clone\_1lb="brain cDNA library"  
/dev\_stage="adult"  
1. .2000  
/gene="hxcT"  
136. .1620  
/gene="hxcT"  
/note="Nat-independent amino acid transporter  
transmembrane protein"  
/codon\_start=1  
/product="cystine/glutamate exchanger"  
/protein\_id="BAB0574.1"

```

/db_xref="GI:13516846"
/translation="MVRKPVSTISKGGYLOGNVNRLPSLGNKEPQGEVKLRKV
TLRGSIIITIGITIGAGTISPKVYQNTGVSMTITWYGVLSLFGALSTAEIGT
TIKRSKGHTYVILEVFGPLPAFVWVWELIIRPATATVSLARGYILPEFIOCEI
PELAKILITAVGIVVWVWVLSMSVWSARIQITFCRLAIIITIVPGVQMLKQOT
ONFKDASGRDSITRPLAFYGMVAVAGWFLNTEVEVEPEKTIPLAICSMAL
VTIGVLTNVAEFTTINAEELLISNAVAFSESLGNFSLAVIFVALSCFGSMNG
VFAVSRLFEVVASREGHLEPILSMIHKRHTPLPAVILHPLTMIIMLSGDSLSLNL
SFARMLFTGLAVAGLIYIRYKCPDMHREPKVLPFLPALSFTCLFMVALSLYSDPFS
GIGFVITLTVGPATYILFTMDKRPWRFRIMSKFSLAMPACDM"

BASE COUNT      505 a      418 c      435 g      642 t
ORIGIN
Alignment Scores:
Pred. No.:      1,3e-29      Length:      2000
Score:          37.00      Matches:      37
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    38.14%      Indels:      0
DB:              9      Gaps:      0

US-09-854-133-586 (1-97) x AB040875 (1-2000)
QY 61 GlusertleucyProPProserProlysgluValThrcysArgGluMetLeuThrGly 80
Db 143 GAAAGCCTGTGTGTGTCACCATCTCCAAAGAGGAGTTACCTGAGGGAATGTTAACGGGA 202
QY 81 GlyCysLeuProTPrPaLaThArGserHisLeuGlyArGArGlyCysSer 97
Db 203 GGCTGCTTCTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCAGC 253

RESULT 11
LOCUS      BC012087      2155 bp      mRNA      linear      PRI 06-AUG-2001
DEFINITION Homo sapiens, Similar to solute carrier family 7, (cationic amino
LOCUS      IMAGE:4562994, mRNA, complete cds.
ACCESSION      BC012087      GI:15082351
VERSION      MGC.
KEYWORDS      Homo sapiens.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE      1 (bases 1 to 2155)
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (02-AUG-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: DCTD/DRP
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
InfoDbgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Buterfield,
Susana Chan, Readman Chiu, Chris Fjell, Erin Gailard, Nan Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McAvery, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNLN at: http://image.lnl.gov
Series: IRAL Plate: 29 Row: 9 Column: 1
This clone was selected for full length sequencing because it

```

```

FEATURES
Source
Location/Qualifiers
1..2155
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20026 IMAGE:4562994"
/tissue_type="Kidney, renal cell adenocarcinoma"
/clone_id="NIH_MGC_14"
/lab_host="DH10B-R"
/ab_vector="potB"
/notes="Vector: potB"
161..1666
/codon_start=1
/product="Similar to solute carrier family 7, (cationic
amino acid transporter, y+ system) member 11"
/protein_id="AAH12087.1"
/db_xref="GI:15082352"
/translation="MVRKPVSTISKGGYLOGNVNRLPSLGNKEPQGEVKLRKV
TLRGSIIITIGITIGAGTISPKVYQNTGVSMTITWYGVLSLFGALSTAEIGT
TIKRSKGHTYVILEVFGPLPAFVWVWELIIRPATATVSLARGYILPEFIOCEI
PELAKILITAVGIVVWVWVLSMSVWSARIQITFCRLAIIITIVPGVQMLKQOT
ONFKDASGRDSITRPLAFYGMVAVAGWFLNTEVEVEPEKTIPLAICSMAL
VTIGVLTNVAEFTTINAEELLISNAVAFSESLGNFSLAVIFVALSCFGSMNG
VFAVSRLFEVVASREGHLEPILSMIHKRHTPLPAVILHPLTMIIMLSGDSLSLNL
SFARMLFTGLAVAGLIYIRYKCPDMHREPKVLPFLPALSFTCLFMVALSLYSDPFS
GIGFVITLTVGPATYILFTMDKRPWRFRIMSKFSLAMPACDM"

BASE COUNT      599 a      443 c      452 g      661 t
ORIGIN
Alignment Scores:
Pred. No.:      1,38e-29      Length:      2155
Score:          37.00      Matches:      37
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    38.14%      Indels:      0
DB:              9      Gaps:      0

US-09-854-133-586 (1-97) x BC012087 (1-2155)
QY 61 GlusertleucyProPProserProlysgluValThrcysArgGluMetLeuThrGly 80
Db 168 GAAAGCCTGTGTGTGTCACCATCTCCAAAGAGAGTTACCTGAGGGAATGTTAACGGGA 227
QY 81 GlyCysLeuProTPrPaLaThArGserHisLeuGlyArGArGlyCysSer 97
Db 228 GGCTGCTTCTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCAGC 278

RESULT 12
LOCUS      AF252872      2482 bp      mRNA      linear      PRI 02-MAY-2001
DEFINITION Homo sapiens cystine/glutamate transporter xCT mRNA, complete cds.
ACCESSION      AF252872
VERSION      AF252872.1 GI:13924719
KEYWORDS      Homo sapiens.
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE      1 (bases 1 to 2482)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE      Chancy,C.D., Kekuda,R., Wang,H., Huang,W., Prasad,P.D., Smith,S.B.
and Ganapathy,V.
JOURNAL      Structure, Function and Regulation of Human Cystine/Glutamate
Transporter in Retinal Pigment Epithelial Cells
Unpublished
2 (bases 1 to 2482)
Wang,H., Prasad,P.D. and Ganapathy,V.
Direct Submission
Submitted (05-APR-2000) Biochemistry & Molecular Biology, Medical
College of Georgia, 1120 15th Street, Augusta, GA 30912, USA
Location/Qualifiers
1..2482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="placenta"

```

```

CDS
232..1737
/codon_start=1
/product="cystine/glutamate transporter xct"
/protein_id="AAK49111.1"
/db_xref="GI:13924720"
/translation="MVRKPVSTISKGYLOGNVNGLRPSICNKEPPGQEVOLKRY.
TLRGVSIILIGIAGIFISPKGVLONGSVSMILTMVCGLSLFGALSYALEGT
TLKSGGHTYILEVGPAPFVRVWELLITPAATAVISLAFGRITLPEFLOCBI
PELAKRLITAVGTYVVMVNSVSMANIQLEITFCGLTALILITVGMOLIKGT
VIGVLTINVAFTTINAEELLISNAVAWFLNFEVENDEKTIPIATISMAI
VEAVRLTIVASREGLPELISIMHVRKTPPLPAVITLPIITIMLPSGDLNLPL
SPARLFTGLAVAGILYLRKCPDMHRRPKVPLFIPALFSFCLFVVALISYSDPST
GIGFVITLGVPAVYLFITWDRKPRFMRISKIRITLQILLEVPEEDKL"

BASE COUNT 735 a 484 c 521 g 742 t
ORIGIN

Alignment Scores:
Pred. No.: 1.56e-29
Score: 37.00 Length: 2482
Percent Similarity: 100.00 Matches: 37
Best Local Similarity: 100.00 Conservative: 0
Query Match: 38.14% Mismatches: 0
Db: 9 Indels: 0
Gaps: 0

US-09-854-133-586 (1-97) x AF252872 (1-2482)

OY 61 GluSerLeuLeuCySPProSerProLySGluValThrcysArgGluMetLeuThrcly 80
DB 239 GAAGGCTGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGAGAAATGTTAACGGGA 298
OY 81 GlyCysLeuProTrrAlaThrArgSerHisLeuGlyArgArgLysCysSer 97
DB 299 GGCTGCTTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCACG 349

RESULT 13
HSA277882 3144 bp mRNA linear PRI 04-JAN-2002
DEFINITION Homo sapiens mRNA for cystine/glutamate transporter (XCT gene).
ACCESSION AJ2277882
VERSION AJ2277882.1 GI:18073361
KEYWORDS cystine/glutamate transporter; XCT gene.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Borsani,G., Manzoni,M., Palacin,M., Plineda,M. and Gasol,E.
Unpublished
2 (bases 1 to 3144)
Bassi,M.T.
TITLE Direct Submission
JOURNAL Submitted (11-MAY-2000) Bassi M.T., Telethon Institute of Genetics
and Medicine, Via Olgettina 58, 20132 Milan, ITALY
FEATURES
source
1..3144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4q28-q32"
/cell_line="NR2, undifferentiated teratocarcinoma cell
1..3128
/feature="XCT"
1..1506
/feature="XCT"
/feature="XCT"
/codon_start=1
/product="cystine/glutamate transporter"
/protein_id="CAC81905.1"
/db_xref="GI:18073362"
/translation="MVRKPVSTISKGYLOGNVNRLSLNKEPPGQEVOLKRY
TLKSGHTYILEVGPAPFVRVWELLITPAATAVISLAFGRITLPEFLOCBI
TIKSGHTYILEVGPAPFVRVWELLITPAATAVISLAFGRITLPEFLOCBI"

CDS
232..1737
/codon_start=1
/product="cystine/glutamate transporter xct"
/protein_id="AAK49111.1"
/db_xref="GI:13924720"
/translation="MVRKPVSTISKGYLOGNVNGLRPSICNKEPPGQEVOLKRY.
TLRGVSIILIGIAGIFISPKGVLONGSVSMILTMVCGLSLFGALSYALEGT
TLKSGGHTYILEVGPAPFVRVWELLITPAATAVISLAFGRITLPEFLOCBI
PELAKRLITAVGTYVVMVNSVSMANIQLEITFCGLTALILITVGMOLIKGT
VIGVLTINVAFTTINAEELLISNAVAWFLNFEVENDEKTIPIATISMAI
VEAVRLTIVASREGLPELISIMHVRKTPPLPAVITLPIITIMLPSGDLNLPL
SPARLFTGLAVAGILYLRKCPDMHRRPKVPLFIPALFSFCLFVVALISYSDPST
GIGFVITLGVPAVYLFITWDRKPRFMRISKIRITLQILLEVPEEDKL"

BASE COUNT 735 a 484 c 521 g 742 t
ORIGIN

Alignment Scores:
Pred. No.: 1.92e-29
Score: 37.00 Length: 3144
Percent Similarity: 100.00 Matches: 37
Best Local Similarity: 100.00 Conservative: 0
Query Match: 38.14% Mismatches: 0
Db: 9 Indels: 0
Gaps: 0

US-09-854-133-586 (1-97) x HSA277882 (1-3144)

OY 61 GluSerLeuLeuCySPProSerProLySGluValThrcysArgGluMetLeuThrcly 80
DB 8 GAAGGCTGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGAGAAATGTTAACGGGA 67
OY 81 GlyCysLeuProTrrAlaThrArgSerHisLeuGlyArgArgLysCysSer 97
DB 68 GGCTGCTTCCCTGGGCAACAAGAGCCACCTGGGAGAGAAAGTGCACG 118

RESULT 14
TCR292491 295 bp DNA linear VRT 16-MAY-2002
DEFINITION Triturus cristatus microsatellite DNA, locus Tcr36.
ACCESSION AJ292491
VERSION AJ292491.1 GI:20798953
KEYWORDS microsatellite; repetitive element.
SOURCE warty newt.
ORGANISM Triturus cristatus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Caudata; Salamandridae; Salamandridae;
Triturus.
REFERENCE
1 Krupa,A.P., Uhle,R., Dawson,D.A., Gentle,L.K., Gibbs,M.,
Aritzen,J.W. and Burke,T.
Microsatellite loci in the crested newt (Triturus cristatus) and
their utility in other newt taxa
Conserv. Genet. 3, 87-89 (2002)
2 (bases 1 to 295)
Krupa,A.P.
TITLE Direct Submission
JOURNAL Submitted (18-JUL-2000) Krupa A.P., Animal and Plant Sciences,
University of Sheffield, Alfred Denny Building, Western Bank,
Sheffield, South Yorkshire, S10 2TN, UNITED KINGDOM
FEATURES
source
1..295
/organism="Triturus cristatus"
/db_xref="taxon:8323"
/clone="TC7363"
/feature="locus Tcr36"
1..295

BASE COUNT 156 a 26 c 69 g 43 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 0.000952
Score: 12.00 Length: 295
Percent Similarity: 100.00 Matches: 12
Best Local Similarity: 100.00 Conservative: 0
Query Match: 12.37% Mismatches: 0
Db: 5 Indels: 0
Gaps: 0

```





[illegible]

```
misc_feature
    /note="repeat match = HSA02250; putative"
    /rpt_family="Alu"
    6390..7080
repeat_region
    /note="ss region (gap in top strand); puta
    complement(6665..6950)
    /note="repeat match = HSA00697; putative"
    /rpt_family="Alu"
    complement(6673..6879)
    /note="repeat match = HSA05100; putative"
    /rpt_family="Alu"
    complement(6884..6967)
    /note="repeat match = HSA06328; putative"
    /rpt_family="Alu"
    complement(7298..7445)
    /note="repeat match = HSA04736; putative"
    /rpt_family="Alu"
    complement(7331..7620)
    /note="repeat match = ALU; putative"
    /rpt_family="Alu"
    complement(7758..7839)
    /note="repeat match = HSA02061; putative"
    /rpt_family="Alu"
    complement(7739..8057)
    /note="repeat match = HSA06256; putative"
    /rpt_family="Alu"
    complement(7837..8068)
    /note="repeat match = HSA00736; putative"
    /rpt_family="Alu"
    9050..9100
misc_feature
    /note="ss region (gap in bottom strand); pu
    complement(9110..9354)
    /note="repeat match = HSA04630; putative"
    /rpt_family="Alu"
    complement(9173..9712)
    /note="repeat match = HSA01827; putative"
    /rpt_family="Alu"
    complement(9677..9723)
    /note="repeat match = HSA03456; putative"
    /rpt_family="Alu"
    complement(9818..9930)
    /note="repeat match = HSA00794; putative"
    /rpt_family="Alu"
    complement(9841..9977)
    /note="repeat match = HSA01882; putative"
    /rpt_family="Alu"
    complement(9873..9989)
    /note="repeat match = HSA05507; putative"
    /rpt_family="Alu"
```

Alignment Scores:	
Pred. No.:	
Score:	0.07
Percent Similarity:	12.00
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	12.37%
	9
	Gaps:
	0

US-09-854-133-586 (1-97) x CH19F15314 (1-41369)

OY	23	ArgLysLysGLuArgLysLysLysArGGLuArgLys	34
Ddb	22456	AGAAAGAAAAGAAAGAAAAAGAGAGAAGAAAG	22421
RESULT	17		
LOCUS	AC101290/c.		
DEFINITION	Mus musculus clone RP23-103D10, LOW-PASS	DNA	linear
ACCESSION	AC101290		HTG 23-NOV-2001
VERSION	AC101290.1		SEQUENCE SAMPLING.
KEYWORDS	HTG; HTGS. PHASEO.		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 68527)  
Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
Mus musculus, clone RP23-103D10  
Unpublished

2 (bases 1 to 68527)  
Nusbaum, C., Lander, E., All, A., Allen, N.,  
Birren, B., Linton, L., Bastien, V., Boguslavsky, L., Bonkhalter, B.,  
Anderson, S., Barina, N., Campiano, A., Chang, J., Chazaro, B.,  
Brown, A., Camarata, J., Collins, S., Collamore, A., Cook, A.,  
Choe, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,  
Cooke, P., Dearlano, K., Dewar, K., Diaz, J., Dodge, S., Ferro, S.,  
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Laroque, K.,  
Lamazares, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,  
McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,  
McLellan, C., Macdonald, P., Major, J., Maylor, J., Nguyen, C.,  
Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., O'Neil, D.,  
O'Brien, C., Norman, C. H., O'Connor, T., O'Donnell, P., Pollard, V.,  
Oliver, J., Peterson, K., Phunhng, P., Plerre, N., Pollard, V.,  
Raymond, C., Reta, R., Riedack, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,  
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zahoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L16367  
Center clone name: 103\_D10

\*\*\*\*\* NOTE: This record contains 88 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 652: contig of 652 bp in length  
\* 653 752: gap of 100 bp  
\* 753 1436: contig of 684 bp in length  
\* 1437 1536: gap of 100 bp  
\* 1537 2226: contig of 690 bp in length  
\* 2227 2326: gap of 100 bp  
\* 2327 2993: contig of 667 bp in length  
\* 2994 3093: gap of 100 bp  
\* 3094 3785: contig of 692 bp in length  
\* 3786 3885: gap of 100 bp  
\* 3886 4575: contig of 690 bp in length  
\* 4576 4675: gap of 100 bp  
\* 4676 5365: contig of 690 bp in length  
\* 5366 5465: gap of 100 bp  
\* 5466 6163: contig of 698 bp in length  
\* 6164 6263: gap of 100 bp  
\* 6264 6962: contig of 699 bp in length  
\* 6963 7062: gap of 100 bp  
\* 7063 7741: contig of 679 bp in length

7742 7841: gap of 100 bp  
\* 7842 8528: contig of 687 bp in length  
\* 8529 8628: gap of 100 bp  
\* 8629 9293: contig of 665 bp in length  
\* 9294 9393: gap of 100 bp  
\* 9394 10062: contig of 669 bp in length  
\* 10063 10162: gap of 100 bp  
\* 10163 10812: contig of 650 bp in length  
\* 10813 10912: gap of 100 bp  
\* 10913 11600: contig of 688 bp in length  
\* 11601 11700: gap of 100 bp  
\* 11701 12394: contig of 694 bp in length  
\* 12395 12494: gap of 100 bp  
\* 12495 13162: contig of 668 bp in length  
\* 13163 13262: gap of 100 bp  
\* 13263 13966: contig of 684 bp in length  
\* 13967 14046: gap of 100 bp  
\* 14047 14738: contig of 692 bp in length  
\* 14739 14838: gap of 100 bp  
\* 14839 15510: contig of 672 bp in length  
\* 15511 15610: gap of 100 bp  
\* 15611 16296: contig of 686 bp in length  
\* 16297 16395: gap of 100 bp  
\* 16397 17057: contig of 661 bp in length  
\* 17058 17157: gap of 100 bp  
\* 17158 17842: contig of 685 bp in length  
\* 17843 17942: gap of 100 bp  
\* 17943 18608: contig of 666 bp in length  
\* 18609 18708: gap of 100 bp  
\* 18709 19410: contig of 702 bp in length  
\* 19411 19510: gap of 100 bp  
\* 19511 20159: contig of 649 bp in length  
\* 20160 20259: gap of 100 bp  
\* 20260 20939: contig of 680 bp in length  
\* 20940 21039: gap of 100 bp  
\* 21040 21719: contig of 680 bp in length  
\* 21720 21819: gap of 100 bp  
\* 21820 222495: contig of 676 bp in length  
\* 22496 22595: gap of 100 bp  
\* 22596 23280: contig of 685 bp in length  
\* 23281 23380: gap of 100 bp  
\* 23381 24065: contig of 685 bp in length  
\* 24066 24165: gap of 100 bp  
\* 24166 24861: contig of 696 bp in length  
\* 24862 24961: gap of 100 bp  
\* 24962 25657: contig of 696 bp in length  
\* 25658 25757: gap of 100 bp  
\* 25758 26424: contig of 667 bp in length  
\* 26425 26524: gap of 100 bp  
\* 26525 27156: contig of 632 bp in length  
\* 27157 27256: gap of 100 bp  
\* 27257 27953: contig of 697 bp in length  
\* 27954 28053: gap of 100 bp  
\* 28054 28751: contig of 698 bp in length  
\* 28752 28851: gap of 100 bp  
\* 28852 29533: contig of 682 bp in length  
\* 29534 29633: gap of 100 bp  
\* 29634 30320: contig of 687 bp in length  
\* 30321 30420: gap of 100 bp  
\* 30421 31064: contig of 644 bp in length  
\* 31065 31164: gap of 100 bp  
\* 31165 31836: contig of 672 bp in length  
\* 31837 31936: gap of 100 bp  
\* 31937 32630: contig of 694 bp in length  
\* 32631 32730: gap of 100 bp  
\* 32731 33394: contig of 664 bp in length  
\* 33395 33494: gap of 100 bp  
\* 33495 34157: contig of 663 bp in length  
\* 34158 34257: gap of 100 bp  
\* 34258 34947: contig of 690 bp in length  
\* 34948 35047: gap of 100 bp  
\* 35048 35724: contig of 677 bp in length  
\* 35725 35824: gap of 100 bp





COMMENT	On Feb 24, 2000 this sequence version replaced gi:5997411
FEATURES	Location/Qualifiers
source	1..76163

BASE COUNT	24100 a	14052 c	14597 g	23414 t
ORIGIN				

Pred. No.:	0.119	Length:	7616
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	9		

05-09-854-133-586 (1-97) x AP000960 (1-76163)

23 ATGLYSLYSGLUARGYSLYSLYSGLUARGLYS 34  
 Db 19555 AGAAAGGAAGAAACAAGAAAAAGAGAGAAAGAAAG 19590

HSJ189G13/c	100167 bp	DNA	linear	PRI 24-FEB-2001
LOCUS				
DEFINITION	Human DNA sequence from clone RP1-189G13 on chromosome 20. Contains an RPL17A (60S ribosomal protein L7A) (SURF3) pseudogene, an RPS4 (40S ribosomal protein S4) pseudogene, ESTs, STSS and GSSs, complete sequence.			
ACCESSION	U170106			

ACCESSION	AL121916
VERSION	AL121916.14
KEYWORDS	HTG; RPL7A; RPS4; SURE3
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1. (Phases 1 to 100167)  
Williams, S.  
Direct Submission  
Submitted (22-FEB-2001) *Current Science*  
100167  
Euteleostomi  
Teleostei  
Catarrhini  
Hominiidae; Homo.

**COMMENT**

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chn20  
IMPORTANT: This sequence is not the entire insert of clone  
Rpl-189g13. It may be shorter because we sequence overlapping  
sections only once, except for a 100 base overlap.  
The true left end of clone Rpl-189g13 is at 1 in this sequence.  
The true left end of clone RP5-1068H6 is at 100068 in this sequence.  
The true right end of clone RP5-1164c1 is at 13278 in this  
sequence. This sequence was finished as follows unless otherwise  
noted: all regions were either double-stranded or sequenced with  
alternate chemistry or covered by high quality data (i.e., phased  
quality >= 30); an attempt was made to resolve all sequencing  
problems, such as compressions and repeats; all regions were  
covered by at least one plasmid subclone or more than one M13  
Rpl-189g13 and the assembly was confirmed by restriction digest.  
Plaster de Jong. For further details see  
http://www.chori.org/naopac/home.htm  
VECTOR: pCIRAC2.

source	Location/Qualifiers
repeat_region	1. .100167
repeat_region	/organism="Homo sapiens"
repeat_region	/db_xref="taxon:9606"
repeat_region	/chromosome="20"
repeat_region	/clone="RP1-189G13"
repeat_region	/clone_id="RPC1-1"
repeat_region	14. .1225
repeat_region	/note="LIM49 repeat: matches 4456. .5781 of consensus"
repeat_region	1236. .1337
repeat_region	/note="46 copies 2 mer tc 80% conserved"
repeat_region	complement(1276. .1378)
repeat_region	/note="match: GSS: Em:A2007965"
repeat_region	1376. .1864
repeat_region	/note="LIM49 repeat: matches 5783. .6294 of consensus"
repeat_region	1987. .2288
repeat_region	/note="AluX repeat: matches 1. .310 of consensus"
repeat_region	2431. .2702
repeat_region	/note="L2 repeat: matches 2241. .2537 of consensus"
repeat_region	3194. .3649
repeat_region	/note="L1PA13 repeat: matches 5702. .6144 of consensus"
repeat_region	3779. .3842
repeat_region	/note="32 copies 2 mer tt 71% conserved"
repeat_region	4654. .4751
repeat_region	/note="MIR repeat: matches 124. .230 of consensus"
repeat_region	4788. .5239
repeat_region	/note="LIM7 repeat: matches 5481. .5907 of consensus"
repeat_region	5240. .5537
repeat_region	/note="AluO repeat: matches 1. .299 of consensus"
repeat_region	5538. .5663
repeat_region	/note="LIMB7 repeat: matches 5363. .5481 of consensus"
repeat_region	5664. .5958
repeat_region	/note="AluX repeat: matches 1. .297 of consensus"
repeat_region	5939. .6048
repeat_region	/note="LIMB7 repeat: matches 5207. .5363 of consensus"
repeat_region	6046. .6269
repeat_region	/note="LIMB7 complement(6138. .6621)"
repeat_region	/note="match: GSS: Em:AQ439122"
repeat_region	6539. .8750
repeat_region	/note="LIM49 repeat: matches 3474. .5706 of consensus"
repeat_region	8751. .9050
repeat_region	/note="AluX repeat: matches 1. .300 of consensus"
repeat_region	9051. .9362
repeat_region	/note="AluX repeat: matches 1. .312 of consensus"
repeat_region	9363. .9762
repeat_region	/note="LIM49 repeat: matches 5706. .6064 of consensus"
repeat_region	9763. .9800
repeat_region	/note="LTR16C repeat: matches 350. .387 of consensus"
repeat_region	complement(9941. .10370)
repeat_region	/note="match: GSS: Em:AQ696486"
repeat_region	10332. .10825
repeat_region	/note="match: GSS: Em:AQ672010"
repeat_region	11276. .11654
repeat_region	/note="LTR16C repeat: matches 1. .366 of consensus"

```

repeat_region 11928..12124
/note="HERVL repeat: matches 4464..4658 of consensus"
repeat_region 12503..13301
/note="HERVL repeat: matches 3127..3940 of consensus"
misc_feature 13588..13691
/note="match: GSS: Em:AQ0548735 Em:AQ0800523
match: SMS: Em:AQ0548735"
repeat_region 13724..13949
/note="LIM4 repeat: matches 3331..3553 of consensus"
repeat_region 13951..14855
/note="LIM4 repeat: matches 5440..6317 of consensus"
repeat_region 14836..16206
/note="L1 repeat: matches 2655..4042 of consensus"
repeat_region 16207..16426
/note="MER3A repeat: matches 1..230 of consensus"
repeat_region 16427..16523
/note="L1 repeat: matches 2560..2655 of consensus"
repeat_region 16531..16740
/note="MER74A repeat: matches 353..558 of consensus"
repeat_region 16746..16805
/note="30 copies 2 mer tg 90% conserved"
repeat_region 16806..17085
/note="MER74A repeat: matches 58..345 of consensus"
repeat_region 17102..17182
/note="MER74A repeat: matches 1..88 of consensus"
repeat_region 17192..17323
/note="LIM4 repeat: matches 3195..3319 of consensus"
repeat_region 17579..17752
/note="MER5A repeat: matches 6..186 of consensus"
misc_feature 17998..18371
/note="match: GSS: Em:B99526"
repeat_region 18871..19048
/note="MIR repeat: matches 85..262 of consensus"
repeat_region 19049..19554
/note="MIR repeat: matches 19..547 of consensus"
repeat_region 19558..19581
/note="12 copies 2 mer tt 100% conserved"
repeat_region 20668..20913
/note="123 copies 2 mer ct 8% conserved"
repeat_region 21007..21238
/note="AluB repeat: matches 2..229 of consensus"
misc_feature 21496..21915
/note="match: GSS: Em:AQ200755"
misc_feature 21510..21716
/note="match: GSS: Em:AQ297835"
repeat_region 21582..21657
/note="Charlie2 repeat: matches 67..144 of consensus"
repeat_region 21808..21901
/note="Charlie2 repeat: matches 254..349 of consensus"
misc_feature 22351..23665
/note="match: GSS: Em:AQ141400"
repeat_region 26991..27304
/note="Charlie2 repeat: matches 2384..3718 of consensus"
repeat_region 27392..27480
/note="LIM4 repeat: matches 2920..3242 of consensus"
repeat_region 27655..27840
/note="MIR repeat: matches 105..189 of consensus"
repeat_region 28167..30593
/note="MIR repeat: matches 71..262 of consensus"
repeat_region 30621..30679
/note="LIM3 repeat: matches 4017..5403 of consensus"
repeat_region 30685..31400
/note="LIM3 repeat: matches 6853..6918 of consensus"
repeat_region 31431..31557
/note="LIM1 repeat: matches 5504..6224 of consensus"
repeat_region 31950..31984
/note="LIM3 repeat: matches 1..130 of consensus"
repeat_region 31980..32672
/note="MIR repeat: matches 513..547 of consensus"
repeat_region 32902..33244
/note="MER21B repeat: matches 96..773 of consensus"
misc_feature 32902..33244
/note="match: GSS: Em:B32572"

```

```

repeat_region 34122..34247
/note="L2 repeat: matches 2572..2710 of consensus"
misc_feature 34327..34694
/note="match: GSS: Em:AQ120726"
misc_feature 34676..35116
/note="match: GSS: Em:B16156"
repeat_region 34720..35038
/note="MIR repeat: matches 226..547 of consensus"
repeat_region 35056..35397
/note="MIR repeat: matches 240..583 of consensus"
repeat_region 35650..35843
/note="MIR repeat: matches 688..890 of consensus"
repeat_region 36180..36250
/note="MIR repeat: matches 1108..1177 of consensus"
repeat_region 36329..36451
/note="LIM1 repeat: matches 6046..6172 of consensus"
repeat_region 36445..36586
/note="LIM1 repeat: matches 5345..5484 of consensus"

```

```

Alignment Scores:
Pred. No.: 0.151 Length: 100167
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.37% Indels: 0
Gaps: 9
DB:

```

US-09-854-133-586 (1-97) x HSI189613 (1-100167)

QY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
 Db 20789 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 20754

#### RESULT 21

AC090033 125267 bp DNA linear MAM 21-JUL-2001  
 LOCUS  
 DEFINITION Felis catus clone RP66-523H23, complete sequence.  
 AC090033  
 VERSION  
 KEYWORDS

AC090033.2 GI:14993725

HTG.  
 SOURCE Felis catus.  
 ORGANISM Felis catus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

#### REFERENCE

AUTHORS

1 (bases 1 to 125267)  
 Aylee, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,  
 Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L.,  
 Grante, S., Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E.,  
 Lee-Lin, S.-O., Legaspi, R., Lim, M., Maduro, O.L., Maduro, V.B.,  
 Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J.,  
 Pearson, R., Prasad, A., Shcherchenko, Y., Snyder, B., Stantiripop, S.,  
 Thomas, J.W., Thomas, P.T., Touchman, J.W., Tsurgou, C., Vogt, J.L.,  
 Walker, M.A., Wetherby, K.D., Zhang, L.-H. and Green, E.D.

NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 125267)

#### TITLE

REFERENCE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

COMMENT

Submitted (21-JUL-2001) NIH Intramural Sequencing Center, 8717  
 Grovmont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 125267)  
 Green, E.D.  
 Direct Submission  
 Submitted (10-FEB-2001) NIH Intramural Sequencing Center, 8717  
 Grovmont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 125267)  
 Green, E.D.  
 Direct Submission  
 Submitted (21-JUL-2001) NIH Intramural Sequencing Center, 8717  
 Grovmont Circle, Gaithersburg, MD 20877, USA  
 On Jul 21, 2001 this sequence version replaced gi:12739798.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>

**CLONE LENGTH:** This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features section.

```
misc_feature
1. .31338
```

BASE COUNT	/note="single clone coverage
37454	a 22462 c 24472 g 40879 t

Tree, NO.:	0.183	Length:	1
Score:	12.00	Matches:	1
Percent Similarity:	100.008	Conservatives:	0
Best Local Similarity:	100.008	Mismatches:	0
Query Match:	12.378	Indels:	0
DB:	4	Gaps:	0

QY 23 ArgLysLysGluArgLysLysArgGluArgLys 34  
 |||||  
 Db 548 AGGAAGAAAGAGAGAAAGAAAGAGAGAAAG 51

LOCUS	125661 bp	DNA	linear	MAM 03-NOV-2007
DEFINITION	Felis catus clone			
ACCESSION	RP86-45908, complete sequence.			
AC087731				

## ORGANISMS

REFERENCE  
AUTHORS  
1 (bases 1 to 125661)  
Ayele, K./ Beckstrom-Sternberg S M  
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.  
1 (bases 1 to 125661)  
Ayele, K./ Beckstrom-Sternberg S M  
Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

Bouffard, G. G., Brinkley, C., Brooks, S., Dietrich, N. L., Granite, S.,  
 Guan, X., Gupta, J., Hadou, S.-L., Idol, J. R., Karlins, E., Lee-Ihn, S.-Q.,  
 Leaspi, R., Lim, M., Maduro, Q. L., Maduro, V. B., Masiello, C.,  
 Meistrin, S. D., McCloskey, J. C., McHowe, J., Pearson, R., Prasad, A.,  
 Tiongsan, K. Y., Snyder, B., Starritrop, S., Thomas, J. W., Thomas, P. J.,  
 Tongson, E. E., Touman, J. W., Tsungien, C., Vogt, J. L., Walker, M. A.,  
 Welgeby, J. D., Zhang, L.-H. and Green, E. D.,  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 125661)  
 Green, E. D.

**TITLE** Direct Submission  
**JOURNAL** Submitted (19-JAN-2001) NIH Intramural Sequencing Center, 8717

REFERENCE 3 (bases 1 to 125661)  
AUTHORS Green P D  
TITLE Salivary IgA, salivary IgG, and salivary IgA:G ratio in patients with periodontitis  
JOURNAL JOURNAL OF PERIODONTOLOGY  
PUBLISHED 1997  
PAGES 101-106  
VOLUME 68  
ISSUE 1  
DOI 10.1902/jper.1997.681101  
CITATION PERIODONTOLOGY 68(1):101-106, 1997

Journal  
Submitted (10-FEB

REFERENCE  
AUTHORS  
4 (bases 1 to 125661)  
Green, E.D.

JOURNAL  
Submitted (03-NOV-  
Grovermont Circo)

On Feb 10, 2001 this sequence version replaced gl:12313759.  
----- Genome Center

Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>

Contact: [nisc\\_mouse@nhgri.nih.gov](mailto:nisc_mouse@nhgri.nih.gov)  
----- Project Information  
Center project name: \_\_\_\_\_

Center project name: BWX  
Center clone name: 459008

This sequence was finished as follows unless otherwise noted all regions were double-stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

**CLONE LENGTH:** This sequence represents the entire insert of this clone unless otherwise noted. If there are overlapping clones, the overlaps are noted in the beginning and end of the Features section.

## FEATURES

misc\_feature

```
misc_feature
47870..47882
/note="clone has a very small overlap with GenBank
Accession Number AC087861 (nucleotides 107321-107365
clone RP86-49M22 (center project name aww)"
47870..47882
```

```
misc_feature    /note="single clone coverage"
                49741.49820
/note="single clone coverage"
```

```
misc_feature /note="single clone coverage"
50031. .50069 /note="single clone coverage"
```

```
misc_feature 67131. 67180
misc_feature /note="single clone coverage"
94316. 125661
```

BASE COUNT	37714	a	24480	c	24041	g	39426	t
ORIGIN								

Alignment Scores:

File: NO.:	0.184	Length:	125661
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	4	Gaps:	0

03 03-834-133-586 (1-97) x AC087731 (1-125661

QY 23 ArgLysLysGLuArgLysLysArgGLuArgLys 34  
 |||||  
 94863 AGCAAGAAAGAGAGAAAGAAAAAAGAGAGAAAGAAG 94828  
 Db



RESULT 23  
AC105404/c 127475 bp DNA linear HTG 21-MAY-2002  
LOCUS Mus musculus chromosome UNK clone RP24-112M13, WORKING DRAFT  
DEFINITION  
SEQUENCE, 6 unordered pieces.  
AC105404  
AC105404.3 GI:21039970  
HTG: HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
KEYWORDS  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
REFERENCE  
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
TITLE 1 (bases 1 to 127475)  
JOURNAL McPherson, J.D. and Waterston, R.H.  
The sequence of Mus musculus clone  
Unpublished  
2 (bases 1 to 127475)  
McPherson, J.D. and Waterston, R.H.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (03-JAN-2002) Genome Sequencing Center, 4444 Forest Park  
JOURNAL Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 127475)  
McPherson, J.D. and Waterston, R.H.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park  
JOURNAL Parkway, St. Louis, MO 63108, USA  
On May 21, 2002 this sequence version replaced g1:18767639.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WTGSC  
Web site: http://genome.wustl.edu/gsc/index.shtml  
Contact: submissions@wustl.wustl.edu  
Project Information  
Center project name: M\_BB011M13

----- Summary Statistics -----  
Sequencing vector: M13, 0%  
Sequencing chemistry: Dye-Primer ET; 0% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 127786 bases at least Q40  
Consensus quality: 128075 bases at least Q30  
Consensus quality: 128255 bases at least Q20  
Insert size: 132000; agarose-fp  
Insert size: 129273; sum-of-ctnigs  
Quality coverage: 14.01 in Q20 bases; agarose-fp  
Quality coverage: 12.18 in Q20 bases; sum-of-ctnigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 2153: contig of 2153 bp in length  
\* 1  
\* 2154 2253: gap of unknown length  
\* 2254 18501: contig of 16248 bp in length  
\* 18502 18601: gap of unknown length  
\* 18602 40805: contig of 22204 bp in length  
\* 40806 40905: gap of unknown length  
\* 40906 81368: contig of 40463 bp in length  
\* 81369 81468: gap of unknown length  
\* 81469 127159: contig of 45691 bp in length  
\* 127160 127259: gap of unknown length  
\* 127260 127475: contig of 216 bp in length.  
Location/Qualifiers  
1. 127475

FEATURES  
source  
/db\_xref="taxon:10090"  
/db\_xref="taxon:10090"

misc\_feature  
1. 2153  
/clone="RP24-112M13"  
/note="assembly\_name:Contig31"  
misc\_feature  
2254. 18501  
/note="assembly\_name:Contig32"  
misc\_feature  
18602. 40805  
/note="assembly\_name:Contig33"  
misc\_feature  
40906. 81368  
/note="assembly\_name:Contig34"  
misc\_feature  
81469. 127159  
/note="assembly\_name:Contig35"  
misc\_feature  
127260. 127475  
/note="assembly\_name:Contig11"  
BASE COUNT 41143 a 23490 c 23116 g 39205 t 521 others  
ORIGIN

Alignment Scores:  
Pred. No.: 0.186 Length: 127475  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC105404 (1-127475)

QY 23 ArgLYSLYSGLARGLYSLYSARGGLARGLYS 34  
DB 1816 ACAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1781

RESULT 24  
AC099249/c 143786 bp DNA linear HTG 12-JUL-2002  
LOCUS Rattus norvegicus clone CH230-191L15, \*\*\* SEQUENCING IN PROGRESS  
DEFINITION  
\*\*\* 84 unordered pieces.  
AC099249  
AC099249.5 GI:21729805  
VERSION  
KEYWORDS HTG: HTGS\_PHASE1.  
SOURCE  
ORGANISM Rattus norvegicus  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

1 (bases 1 to 143786)  
Muzny, D.M., Adams, C., Adio-Oduola, B., All-rosman, F.R., Allen, C.,  
Alsbrooks, S.L., Amaral-Lungue, H.C., Are, J.R., Ayele, M., Banks, T.,  
Barbata, J., Benton, J., Bimaye, K., Blankenburg, K., Bonnin, D.,  
Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,  
Buhay, C., Burck, P., Burkett, C., Burrell, R.L., Byrd, N.C.,  
Carroll, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,  
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,  
Cleveland, C.D., Cox, C., Coyle, M.D., Dethorne, S.R., David, R.,  
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,  
Delaney, K.R., Delgado, O., Denna, A.L., Ding, Y., Dinh, H.H.,  
Douthett, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,  
Einhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escoto, M.,  
Farrar, C., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,  
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,  
Gorelli, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hogue, M., Holloway, C., Hollins, L.,  
Homs, F., Howard, S., Huber, J., Hulik, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,  
Karissom, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtenberg, O., Liu, C., Liu, J., Liu, W., Loussed, H.,  
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Manshwar, M., Mapua, P., Martin, R., Martindale, A., Martinez, E.,  
Massey, E., Mawlin, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,  
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,  
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogah, M., Okunolu, G.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Oragunye, N., Oriado, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L., Oulles, M., Ren, Y.,  
Rivers, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Severy, G.,  
Scherer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,  
Sodergren, E., Soneike, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Swatek, L., Taylor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Wang, S.,  
Uman, K., Vasquez, L., Vera, V., Villalón, D., Vlnson, R., Wang, S.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wlarczyk, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G., and Gibbs, R.  
Direct Submission  
Unpublished  
2 (bases 1 to 143786)  
Worley, K.C.  
Direct Submission  
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 143786)  
Worley, K.C.  
Direct Submission  
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 11, 2002 this sequence version replaced g1:18846141.

Center: Baylor College of Medicine

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GUS

Center clone name: CH230-191115

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 49247 bases at least Q40

Consensus quality: 52098 bases at least Q30

Consensus quality: 53978 bases at least Q20

NOTE: Estimated insert size may differ from sequence length

(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).

NOTE: This is a 'working draft' sequence. It currently

consists of 84 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1284: contig of 1284 bp in length  
1285: gap of unknown length  
1385: contig of 1198 bp in length  
2682: gap of unknown length  
3710: contig of 1028 bp in length  
3711: gap of unknown length  
3811: contig of 1622 bp in length  
5432: gap of unknown length  
5532: contig of 1464 bp in length  
6996: gap of unknown length  
7097: gap of unknown length  
8254: contig of 1158 bp in length  
8354: gap of unknown length  
9457: contig of 1103 bp in length  
9557: gap of unknown length  
10583: contig of 1026 bp in length  
10584: gap of unknown length  
11703: contig of 1020 bp in length  
11704: gap of unknown length  
11803: contig of 1004 bp in length  
11804: gap of unknown length  
12907: gap of unknown length

12908: contig of 1006 bp in length  
13914: gap of unknown length  
14013: contig of 1068 bp in length  
14014: gap of unknown length  
15082: contig of 1383 bp in length  
15182: gap of unknown length  
15764: contig of 1018 bp in length  
16765: gap of unknown length  
17883: contig of 1069 bp in length  
17983: gap of unknown length  
19052: contig of 1639 bp in length  
19152: gap of unknown length  
20791: contig of 1017 bp in length  
20891: gap of unknown length  
21907: contig of 1043 bp in length  
22008: gap of unknown length  
23051: contig of 1609 bp in length  
23151: gap of unknown length  
24760: contig of 1387 bp in length  
24860: gap of unknown length  
26247: contig of 1296 bp in length  
26347: gap of unknown length  
27642: contig of 1296 bp in length  
27742: gap of unknown length  
27743: contig of 1611 bp in length  
29353: gap of unknown length  
29453: contig of 1478 bp in length  
30931: gap of unknown length  
31031: contig of 1069 bp in length  
31032: gap of unknown length  
32101: contig of 1854 bp in length  
32201: gap of unknown length  
34055: contig of 1622 bp in length  
34155: gap of unknown length  
35777: contig of 1454 bp in length  
35877: gap of unknown length  
37331: contig of 1527 bp in length  
37431: gap of unknown length  
38958: contig of 1595 bp in length  
39058: gap of unknown length  
40553: contig of 1357 bp in length  
40752: gap of unknown length  
42109: contig of 1879 bp in length  
42209: gap of unknown length  
44088: contig of 1328 bp in length  
44189: gap of unknown length  
44899: contig of 1429 bp in length  
45517: gap of unknown length  
45617: contig of 1429 bp in length  
47046: gap of unknown length  
47145: contig of 1831 bp in length  
48976: gap of unknown length  
49076: contig of 1744 bp in length  
49077: gap of unknown length  
49777: contig of 1365 bp in length  
50821: gap of unknown length  
50920: contig of 1065 bp in length  
52285: gap of unknown length  
52385: contig of 1065 bp in length  
53450: gap of unknown length  
53450: contig of 1312 bp in length  
54862: gap of unknown length  
54962: contig of 1238 bp in length  
56200: gap of unknown length  
56300: contig of 1521 bp in length  
57821: gap of unknown length  
57921: contig of 1007 bp in length  
58928: gap of unknown length  
59028: contig of 1177 bp in length  
60205: gap of unknown length  
60306: contig of 1228 bp in length  
61533: gap of unknown length  
61633: contig of 1176 bp in length  
62810: gap of unknown length  
62910: contig of 1051 bp in length  
63960: gap of unknown length  
64060: contig of 1503 bp in length  
65563: gap of unknown length  
65663: contig of 1702 bp in length

```
misc_feature misc_feature misc_feature misc_feature misc_feature misc_feature misc_feature misc_source Location/Qualifiers
```

---

```
1..145722 /organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="RP11-21H2"  
/clone_id="RPCI-11 Human Male BAC"  
1..10503  
/note="assembly_fragment  
clone_end:sp6  
vector_side:left"  
10604..12771  
/note="assembly_fragment"  
12872..14989  
/note="assembly_fragment"  
15090..16911  
/note="assembly_fragment"  
17012..19684  
/note="assembly_fragment"  
19785..22088  
/note="assembly_fragment"  
22189..25445  
/note="assembly_fragment"
```

	misc_feature	25546..30628	/note="assembly-fragment"	
	misc_feature	30729..37323	/note="assembly-fragment"	
	misc_feature	37424..42575	/note="assembly-fragment"	
	misc_feature	42676..79708	/note="assembly-fragment"	
	misc_feature	79809..86724	/note="assembly-fragment"	
	misc_feature	86825..109636	/note="assembly-fragment"	
	misc_feature	109737..139934	/note="assembly-fragment"	
	misc_feature	140055..145722	/note="assembly-fragment"	
	misc_end:rv			
	clone_slide:right"			
BASE COUNT	44952	a	27253	c
ORIGIN			45325	t
				1401 others

Alignment Scores:	
Pred. No.:	0 209
Score:	12.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	12.37%
DB:	2
	Gaps:
	0
	Length:
	145722
	Matches:
	12
	Conservative:
	0
	Mismatches:
	0
	Indels:
	0
	Gaps:
	0

US-09-854-133-586 (1-97) x AC015503 (1-145722),

QY	23	ArgLysLysGluArgLysLysArgGluArgLys	34
Db 104218	AGAAAGAAAGAGAGAAAAAGAAAGAGAAAG	104255	

RESULT 26					
AC004932/c					
LOCUS					
DEFINITION	151183 bp	DNA	linear		
ACCESSION	Homo sapiens clone RP5-943F2, complete sequence.			PRI 25-MAR-2001	
VERSION	AC004932				

VERSION	ACU04932.4	GI:13446338
KEYWORDS	HTG.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthelia; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 151183)  
Waterston P H

TITLE	The sequence of Homo sapiens clone
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 151183)
AUTHORS	Waterston, P H

**TITLE** Direct Submission  
**JOURNAL** Submitted (12-JUN-1998) Genome Sequencing Center, Washington  
University School of Medicine, 4440 Forest Park Parkway, St. Louis  
MO 63108, USA  
**REFERENCE** 3 (Bases 1 to 151193)

**AUTHORS** (names & full addresses)  
**TITLE** Waterston, R.H.  
**Direct Submission**  
**Submitted** (07-JUL-2000) **Genome Sequencing Center**

**REFERENCE**  
1. **AUTHORS**  
University School of Medicine, 4444 Forest Park Parkway, St. Louis  
MO 63108, USA  
4 (bases 1 to 151183)  
Waterston, R.

**REFERENCE**  
Submitted (30-SEP-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
5 (bases 1 to 151183)

AUTHOR'S  
 TITLE  
 WATERSTON, R.H.  
 Direct Submission  
 Submitted (25-MAR-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis  
 MO 63108, USA

COMMENT	On Mar 25, 2001 this sequence version replaced g1:8954173
FEATURES	Center project name: H.DJ09430702.
source	Location/Qualifiers
	1..151183

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP5-943F2"
BASE COUNT      42404 a 34087 c 33409 g 41283 t
ORIGIN

```

Alignment Scores:	
Pred. No.:	
Score:	0.216
Percent Similarity:	151183
Best Local Similarity:	12.00
Query Match:	100.00%
AB:	100.00%
	12.37%
Length:	151183
Matches:	12
Conservative:	0
Mismatches:	0
Indels:	0
	0

gaps:  
US-09-854-133-586 (1-97) x AC004932 (1-151183)

Accession	Protein	Length
23	ArgLysIysSGIuaqGLysLysLysArgGIuaqGLys	34
31832	AGAAAAAAGAGACAAAAAGAAAAAGAAAAAGAAAA	31797

ACCESSION	DEFINITION	LOCUS	RESULTS
AC011275	Homo sapiens clone RP11-13J15, WORKING DRAFT SEQUENCE, 7 unordered pieces.	152959 bp DNA	1 linear HMG 20-SEP-2000
AC011276			

VERSION	AC011275.3	GI:10198431
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens	

REFERENCE  
AUTHORS  
Birren, B. (1963) "A  
Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
Mammalia; Euteria; Primates; Carnivora; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 152959)

TITLE	Homo sapiens, clone RP1-13715
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 152959)
AUTHORS	Birren, B., Clinton, I., Nussbaum, C., and Lander, E.

Baldwin, J., Barina, N., Beckert, J., Boguslavsky, L., Boukhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Dealliano, K., Dewar, K., Ferreira, P., Fitzhugh, W., Domino, M., Donelan, J., Doyle, M., Lander, E., Allen, N., Anderson, M., Balwin, J., Barina, N., Beckert, J., Boguslavsky, L., Boukhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A., Cooke, P., Dealliano, K., Dewar, K., Ferreira, P., Fitzhugh, W., Domino, M., Donelan, J., Doyle, M.

Galgagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczký, J., Lien, C., Locke, K., Macdonald, P., Marquis, N., McEwan, P., McGuire, A., Mervin, J., Morris, D., O'Connell, W., Forrest, C., Funke, R., Gage, D.,

Morrow, J., Naylor, J., Norman, C. H., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testave, S., Tirrell, D., Vaccaro, J., Wacker, M., McLaughlin, J., Meldrum, J., O'Connor, T., O'Donnell, P.

TITLE  
JOURNAL  
Myman, D., Ye, W. J., Zimmer, A. and Zody, M., Wheeler, J., Wu, X.,  
Direct Submission  
Submitted (05-Oct-1999) Whitehead Institute/MIT Center for Genomic  
Research, 320 Charles Street, Cambridge, MA 02139, USA

On Sep 20, 2000 this sequence version replaced gi:1930771... All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) <http://ftp.genome.washington.edu/rm/>

-----  
Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www.wi.mit.edu>

```

Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: I3338
Center clone name: 13 T 15

```

```
----- Summary Statistics -----
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
```

AC011492/c

Anderson, S., Barlow, C., Burns, M., Brown, A.,  
Burkett, G., Castle, A.,  
Douglas, J., Boukhalter, B.,

Anderson, S., Balwin, C., Barmann, R., Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G., Castle, A.,





Consensus quality: 163523 bases at least Q20  
Insert size: 148000, 220000





REFERENCE  
AUTHORS

Rattus.  
1 (bases 1 to 171506)

Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralungu,H.C., Are,J.R., Ayele,M., Banks,T., Barbata,J., Benton,J., Blinage,K., Blankenburg,K., Bonnin,D., Buhay,C., Burch,S., Bivela,M., Brown,E., Brown,M., Bryant,N.P., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earmhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B., Homai,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Krtovic,J., Kuresh,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lileu,C., Liu,J., Liu,W., Louleaged,H., Lorado,R.U., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Mohabbat,K., Morgan,M., Morris,S., Nguyen,N., Nickerson,E., Nwokkenwo,S., Oguh,M., Nguyen,N., Oragunye,N., Oyedelo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shoostran,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmari,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNALREFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

Unpublished  
2 (bases 1 to 171506)  
Worley,K.C.  
Submitted (06-May-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 171506)  
Worley,K.C.  
Submitted (14-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Jul 14, 2002 this sequence version replaced gi:20799958.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GVZP  
Center clone name: CH230-313G4  
----- Summary Statistics  
Sequencing vector: Plasmid;  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 110613 bases at least Q40  
Consensus quality: 116745 bases at least Q30  
Consensus quality: 121535 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently consists of 64 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

1	1085	contig of 1085 bp in length
*	1185	gap of unknown length
*	1186	contig of 1045 bp in length
*	2231	gap of unknown length
*	2331	contig of 1442 bp in length
*	3773	gap of unknown length
*	3873	contig of 1068 bp in length
*	4941	gap of unknown length
*	5041	contig of 1040 bp in length
*	6081	gap of unknown length
*	6181	contig of 1404 bp in length
*	7585	gap of unknown length
*	7685	contig of 1261 bp in length
*	8946	gap of unknown length
*	9046	contig of 1621 bp in length
*	10667	gap of unknown length
*	10667	contig of 1347 bp in length
*	12114	gap of unknown length
*	12214	contig of 1563 bp in length
*	13777	gap of unknown length
*	13877	contig of 1072 bp in length
*	14949	gap of unknown length
*	15048	contig of 1245 bp in length
*	15294	gap of unknown length
*	16293	contig of 1368 bp in length
*	16394	gap of unknown length
*	17661	contig of 1896 bp in length
*	17662	gap of unknown length
*	17862	contig of 1291 bp in length
*	19758	gap of unknown length
*	19857	contig of 1453 bp in length
*	21149	gap of unknown length
*	21248	contig of 1453 bp in length
*	22701	gap of unknown length
*	22702	contig of 1453 bp in length
*	22802	gap of unknown length
*	24254	contig of 1453 bp in length
*	24355	gap of unknown length
*	25665	contig of 1310 bp in length
*	25764	gap of unknown length
*	27300	contig of 1536 bp in length
*	27301	gap of unknown length
*	29535	contig of 2135 bp in length
*	29536	gap of unknown length
*	30916	contig of 1281 bp in length
*	30917	gap of unknown length
*	31017	contig of 1501 bp in length
*	32518	gap of unknown length
*	32618	contig of 1266 bp in length
*	33884	gap of unknown length
*	33983	contig of 1905 bp in length
*	35888	gap of unknown length
*	35889	contig of 1165 bp in length
*	37154	gap of unknown length
*	37254	contig of 2041 bp in length
*	39295	gap of unknown length
*	40996	contig of 1602 bp in length
*	41096	gap of unknown length
*	43689	contig of 2593 bp in length
*	43789	gap of unknown length
*	46865	contig of 3076 bp in length
*	46965	gap of unknown length
*	49295	contig of 2330 bp in length
*	49396	gap of unknown length
*	51847	contig of 2452 bp in length
*	51948	gap of unknown length
*	54027	contig of 2080 bp in length



misc_feature	1. .1103	/note="assembly-fragment"
misc_feature	1204. .3170	/note="assembly-fragment"
misc_feature	3271. .4837	/note="assembly-fragment"
misc_feature	4938. .7078	/note="assembly-fragment"
misc_feature	7179. .9746	/note="assembly-fragment"
misc_feature	9047. .12702	/note="assembly-fragment"
misc_feature	12803. .14951	/note="assembly-fragment"
misc_feature	15052. .17670	/note="assembly-fragment"
misc_feature	17771. .21265	/note="assembly-fragment"
misc_feature	21366. .25812	/note="assembly-fragment"
misc_feature	25913. .30054	/note="assembly-fragment"
misc_feature	30155. .34939	/note="assembly-fragment"

BASE COUNT	vector_side:right	others
53761 a	31812 c	52614 t
ORIGIN	31678 g	2706 others

1

-133-586	(1-97)	x AC027484	(1-172571)
ArgpLysGluArgLysLysLysArgGluArgLys 34			
AGAAAGAAAGAAAGAAAGAAAGACAGAAAGAG 127696			
AC116823	Mus musculus clone RP24-365C10,	173700 bp	DNA
			linear HTG 12-JUL-2002
pieces.			HTG 12-JUL-2002
AC116823	GI:21735055		7 ordered
HTG:	HTGS_PHASE2,	HTGS_DRAFT,	HTGS_FULLTOP.
house mouse,			
Mus musculus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Murinae; Mus.			
1 (bases 1 to 173700)			
Birten,B., Nusbäum,C. and Lander,E.			
Unpublished			
2 (bases 1 to 173700)			
Birten,B., Linton,L., Nusbäum,C., Lander,E., All,A., Allen,N.,			
Anderson,S., Barina,N., Bastien,V., Bloom,T., Boguslavsky,L.,			
Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J.,			
Chazaro,B., Chopel,Y., Colangelo,M., Collins,S., Collamore,A.,			
Cook,A., Cooke,P., Deatrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,			

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Hago, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K., Lamazares, R., Landers, T., Lehocsky, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Mathews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 173700)

Birren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A., Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Mathews, C., McCarthy, M., Meldrum, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Direct Submission  
Submitted (12-JUL-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 12, 2002 this sequence version replaced gi:21700702.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L25576

Center clone name: 365\_C-10

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 172773 bases at least Q40

Consensus quality: 172932 bases at least Q20

Insert size: 172000; agarose-fp

Quality coverage: 8.9 in Q20 bases; sum-of-contigs

Quality coverage: 8.8 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 7 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced

\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 612: contig of 612 bp in length  
\* 613 712: gap of 100 bp  
\* 713 2720: contig of 2008 bp in length  
\* 2721 2820: gap of 100 bp  
\* 2821 7299: contig of 4479 bp in length  
\* 7300 7399: gap of 100 bp  
\* 7400 15732: contig of 8333 bp in length  
\* 15733 15832: gap of 100 bp  
\* 15833 27629: contig of 11797 bp in length  
\* 27630 27729: gap of 100 bp  
\* 27730 67118: contig of 39389 bp in length  
\* 67119 67218: gap of 100 bp  
\* 67219 173700: contig of 106482 bp in length.  
Location/Qualifiers  
1. 173700  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="RP24-365C10"  
/clone\_id="RPCI-24 Male Mouse BAC"

FEATURES  
source  
misc\_feature  
1..612  
/note="assembly-fragment"  
misc\_feature  
713..2720  
/note="assembly-fragment"  
misc\_feature  
2821..7299  
/note="assembly-fragment"  
misc\_feature  
7400..15732  
/note="assembly-fragment"  
misc\_feature  
15833..27629  
/note="assembly-fragment"  
misc\_feature  
27730..67118  
/note="assembly-fragment"  
misc\_feature  
67219..173700  
/note="assembly-fragment"

Alignment Scores:  
Pred. No.: 0.243 Length: 173700  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC116823 (1-173700)

QY 23 ArglySLySGlUArGLySLySArgGLUArGLyS 34

Db 571 AGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 536

RESULT 36 AC027063 173735 bp DNA linear HTG 26-MAY-2000

LOCUS AC027063 Homo sapiens clone RP11-605A10, WORKING DRAFT SEQUENCE, 21

DEFINITION Unordered pieces.

ACCESSION AC027063 GI:8077002

VERSION AC027063.3 GI:8077002

HTG: HTGS\_PHASE1; HTGS\_DRAFT.

KEYWORDS Homo sapiens.

SOURCE Homo sapiens

ORGANISM Homo sapiens

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 173735)

Homo sapiens, Clone RP11-605A10

2 (bases 1 to 173735)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,

Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,

Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,

Campanaro, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,

TITLE  
JOURNAL

## COMMENT

Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardy, S., Glade, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karatas, A.,  
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Liu, G., Locke, K., MacDonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,  
Murphy, T., Naylor, T. M., Norman, C. H., O'Connor, T., O'Donnell, P.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Talamas, J.,  
Strange-Thomson, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessier, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Submitted (26-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 25, 2000 this sequence version replaced gi:7637290.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

Project Information  
Center project name: L7466

Center clone name: 605\_A\_10

Summary Statistics  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 163013 bases at least Q40  
Consensus quality: 168440 bases at least Q30  
Consensus quality: 170590 bases at least Q20  
Insert size: 17000; agarose-fp  
Insert size: 171735; sum-of-ctnigs  
Quality coverage: 4.6 in Q20 bases; agarose-fp  
Quality coverage: 4.5 in Q20 bases; sum-of-ctnigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 21 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1198: contig of 1198 bp in length  
\* 1199 1298: gap of 100 bp  
\* 1299 2804: contig of 1506 bp in length  
\* 2805 2904: gap of 100 bp  
\* 2905 4901: contig of 1997 bp in length  
\* 4902 5001: gap of 100 bp  
\* 5002 8342: contig of 3341 bp in length  
\* 8343 8442: gap of 100 bp  
\* 8443 11290: contig of 2848 bp in length  
\* 11291 11390: gap of 100 bp  
\* 11391 15286: contig of 3896 bp in length  
\* 15287 15386: gap of 100 bp  
\* 15387 19353: contig of 3867 bp in length  
\* 19354 19453: gap of 100 bp  
\* 19454 24735: contig of 5282 bp in length  
\* 24736 24835: gap of 100 bp  
\* 24836 30442: contig of 5607 bp in length  
\* 30443 30542: gap of 100 bp  
\* 30543 38141: contig of 7599 bp in length  
\* 38142 38241: gap of 100 bp  
\* 38242 45256: contig of 7015 bp in length

## FEATURES

## SOURCE

45257 45356: gap of 100 bp  
\* 45357 52387: contig of 7031 bp in length  
\* 52388 52487: gap of 100 bp  
\* 52488 60034: contig of 7547 bp in length  
\* 60035 60134: gap of 100 bp  
\* 60135 68588: contig of 8454 bp in length  
\* 68589 68688: gap of 100 bp  
\* 68689 76018: contig of 7330 bp in length  
\* 76019 76118: gap of 100 bp  
\* 76119 86511: contig of 10393 bp in length  
\* 86512 86611: gap of 100 bp  
\* 86612 95770: contig of 9158 bp in length  
\* 95771 95870: gap of 100 bp  
\* 95871 107153: contig of 11283 bp in length  
\* 107154 107253: gap of 100 bp  
\* 107254 125495: contig of 18242 bp in length  
\* 125496 125595: gap of 100 bp  
\* 125596 144262: contig of 18667 bp in length  
\* 144263 144363: gap of 100 bp  
\* 144363 173735: contig of 29373 bp in length.

Location/Qualifiers

1. 173735  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="RP11-605A10"  
/clone\_lib="RP11-605A10 Human Male BAC"  
1 1198  
/note="assembly-fragment"  
1299 2804  
/note="assembly-fragment"  
2905 4901  
/note="assembly-fragment"  
5002 8342  
/note="assembly-fragment"  
8443 11290  
/note="assembly-fragment"  
11391 15286  
/note="assembly-fragment"  
15387 19353  
/note="assembly-fragment"  
19454 24735  
/note="assembly-fragment"  
24836 30442  
/note="assembly-fragment"  
30543 38141  
/note="assembly-fragment"  
38242 45256  
/note="assembly-fragment"  
45357 52387  
/note="assembly-fragment"  
52488 60034  
/note="assembly-fragment"  
60135 68588  
/note="assembly-fragment"  
68689 76018  
/note="assembly-fragment"  
76119 86511  
/note="assembly-fragment"  
86612 95770  
/note="assembly-fragment"  
95871 107153  
/note="assembly-fragment"  
107254 125495  
/note="assembly-fragment"  
125596 144262  
/note="assembly-fragment"  
144363 173735  
vector\_side:right  
144363 173735  
/note="assembly-fragment"

misc\_feature 52941 a 35819 c 34775 g 48197 t 2003 others  
BASE COUNT  
ORIGIN

Alignment Scores:

Pred. No.:	0.244	Length:	173735
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x AC027063 (1-173735)

QY 8 HisalaseienglyaspsercluthtleuSergin 19  
 |||||  
 Db 159052 CATGCCAGCCTGGGACAGTGAAGACCTGCTCA 159087

RESULT 37  
 AC105990/c  
 LOCUS  
 DEFINITION Mus musculus clone RP24-175E7, WORKING DRAFT SEQUENCE, 4 ordered pieces.  
 AC105990  
 VERSION AC105990.4 GI:22381922  
 HTG: HTGS\_PHASE2; HTGS\_DRAFT; HTGS\_FULLTOP.  
 KEYWORDS house mouse.  
 SOURCE Mus musculus  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 175398)  
 AUTHORS Birren, B., Nusbaum, C. and Lander, E.  
 TITLE Mus musculus, clone RP24-175E7  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 175398)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Batra, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gande, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hades, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 DIRECT SUBMISSION  
 Submitted (11-JAN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 175398)  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

Alignment Scores:

Pred. No.:	0.246	Length:	175398
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x AC105990 (1-175398)

QY 23 ArglyslgslunarglyslslyslsArgGLuArgLys 34

Viell, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
 DIRECT SUBMISSION  
 Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Aug 21, 2002 this sequence version replaced g1:20149427.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
 Project Information  
 Center project name: L20237  
 Center clone name: 175\_E-7  
 Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 174895 bases at least Q40  
 Consensus quality: 174990 bases at least Q30  
 Consensus quality: 170000; agarose-1p  
 Insert size: 175098; sum-of-contigs  
 Quality coverage: 12.5 in Q20 bases; agarose-1p  
 Quality coverage: 12.1 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.  
 This sequence will be replaced  
 by the finished sequence as soon as it is available and the accession number will be preserved.  
 1 17607: contig of 17607 bp in length  
 \* 17608 17707: gap of 100 bp  
 \* 17708 42885: contig of 25178 bp in length  
 \* 42886 42985: gap of 100 bp  
 \* 42986 66633: contig of 23648 bp in length  
 \* 66634 66733: gap of 100 bp  
 \* 66734 175398: contig of 10865 bp in length.  
 Location/Qualifiers  
 1. 175398  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="RP24-175E7"  
 /clone-lib="RPCT-24 Male Mouse BAC"  
 1. 17607  
 /note="assembly-fragment"  
 misc\_feature  
 17708..42885  
 /note="assembly-fragment"  
 misc\_feature  
 42986..66633  
 /note="assembly-fragment"  
 misc\_feature  
 66734..175398  
 /note="assembly-fragment"  
 BASE COUNT 54842 a 32492 c 32477 g 55287 t 300 others  
 ORIGIN







```

Alignment Scores:
Pred. No.:      0.249      Length:      178361
Score:          12.00      Matches:      12
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches:  0
Query Match:    12.37%      Indels:      0
DB:             9           Gaps:        0

US-09-854-133-586 (1-97) x AC006042 (1-178361)

```

RESULT	41				
AC107305					
LOCUS					
DEFINITION		178733 bp	DNA	1 linear	HTG_23-AUG-2002
ACCESSION	AC107305	Homio sapiens chromosome 12 clone RP11-605A10,			WORKING DRAFT
VERSION	AC107305	SEQUENCE, 1 ordered piece.			
KEYWORDS	AC107305.11	GI:22380672			
SOURCE	HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_ACTIVEFIN.				
ORGANISM	human.				
	Homio sapiens				

REFERENCE  
AUTHORS

Mammalia; Euteleostomi; Osteichthyes; Choroiceti; Catarrhini; Vertebrata; Euteleostomi;  
1 (Bases 1 to 178733)

Muzny,D.M., Adams,C., Adlo-oduola,B., Ali-osman,F.R., Allen,C.,  
Alsbrooks,S.L., Amaratunga,H.C., Are,J.R., Ayale,M., Banks,T.,  
Barbaris,J., Benton,J., Blincoe,K., Blankenburg,K., Bonini,D.,  
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,  
Buhy,C., Burch,P., Burnett,C., Butrell,K.L., Byrd,N.C.,  
Carron,T.F., Carter,M., Cavazos,S.R., Checko,J., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Dayla,M.L., Davis,C., Davy-Carroll,L., Dedertch,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Donnhaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Eannhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,  
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,  
Gabisi,A., Gao,Y., Garcia,A., Garner,T., Garza,N., Gill,R.,  
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,  
Harris,C., Harris,C., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,  
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,  
Homsl,F., Howard,S., Huber,J., Hulyk,S., Hunne,T., Jackson,L.E.,  
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvach,J., Kovar,C.,  
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,

**COMMENT**

```

Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Drafting Center Code: MMR
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: HDV
Center clone name: RP11-605A10
----- Summary Statistics -----
Sequencing vector: plasmid;
Sequencing vector: M13;
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 178630 bases at least Q40
Consensus quality: 178699 bases at least Q30
Consensus quality: 178729 bases at least Q20
Estimated insert size: 185737; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html),
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
  I 178733: contig of 178733 bp in length.
Location/Qualifiers
  I .178733

```

```

BASE COUNT      53253 a 37213 c 36488 g 51778 t
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /chromosome="12"
                  /clone="RP11-605A10"
                  1 other:

```

```

ORIGIN
Alignment Scores:
Pred. No.: 0.25 Length: 178733
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.37% Indels: 0
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC107305 (1-178733)
QY 8 HisAlaSerLeuGlyAspSerGluThrLeuSerGln 19
|||||
Db 48057 CATGCCAGCTGGGGACAGTGAACCTGTCTCA 48092

RESULT 42
AC121904 179343 bp DNA linear HTG 21-MAY-2002
LOCUS Mus musculus chromosome UNK clone RP24-17507, WORKING DRAFT
DEFINITION
AC121904.1 GI:21040026
AC121904.1 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS Mus musculus.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 179343)
McPherson,J.D. and Waterston,R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 179343)
McPherson,J.D. and Waterston,R.H.
Direct Submission
Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
Project Information
-----
Center project name: M.BB0175007

----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: Plasmid; 100%
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 175055 bases at least Q40
Consensus quality: 175869 bases at least Q30
Consensus quality: 176611 bases at least Q20
Insert size: 157000; agarose-fp
Insert size: 179794; sum-of-contigs
Quality coverage: 0.00 in Q20 bases; agarose-fp
Quality coverage: 13.00 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently
consists of 7 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 4236: contig of 4236 bp in length
* 4237 4336: gap of unknown length
* 4337 10576: contig of 6240 bp in length
* 10577 10676: gap of unknown length
* 10677 31239: contig of 20563 bp in length

```

```

FEATURES
source
1..179343
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP24-17507"
1..4236
/feature="assembly_name:Contig28"
4337..10576
/feature="assembly_name:Contig29"
10677..31239
/feature="assembly_name:Contig30"
31340..49523
/feature="assembly_name:Contig31"
49624..82058
/feature="assembly_name:Contig32"
82159..114218
/feature="assembly_name:Contig33"
114319..179343
/feature="assembly_name:Contig34"
/feature="assembly_name:Contig34"

BASE COUNT 55906 a 34154 c 33514 g 55156 t 613 others

ORIGIN
Alignment Scores:
Pred. No.: 0.25 Length: 179343
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.37% Indels: 0
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x AC121904 (1-179343)
QY 23 ArgIyLysGluArgGlyLysArgGluArgLys 34
|||||
Db 4159 AGAAGAAAGAGAGAGAGAGAGAGAGAGAG 4124

RESULT 43
AC108976 180336 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-28717, *** SEQUENCING IN PROGRESS
DEFINITION
AC108976
AC108976.3 GI:21737546
AC108976.3 HTG: HTGS_PHASE1.
VERSION HTG: HTGS_PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 180336)
Muzny,D.M., Adams,C., Adio-Oduola,B., All-soman,F.R., Allen,C.,
Aisbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayala,M., Banks,T.,
Bardaric,J., Benton,J., Blinaga,K., Blumenthal,K., Bonnin,D.,
Bouck,J., Bowler,S., Briceva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Buttrill,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davis,M., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabstl,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,

```

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

COMMENT

Correll, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,  
Harris, C., Harris, K., Hart, M., Hayak, P., Hawes, A., Hernandez, J.,  
Hernandez, O., Hodgson, A., Hoggues, M., Holloway, C., Hollins, B.,  
Homsl, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,  
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudan, S.,  
Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,  
Kratovich, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,  
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louised, H.,  
Lopez, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,  
Mansueti, M., Mapua, P., Martin, R., Meador, M., Mel, G., Martinez, E.,  
Massey, E., Mawhney, E., McLeod, M.P., Morgan, M., Morris, S.,  
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Nguyen, M., Nguyen, N.,  
Moser, M., Neal, D., Newton, J., Newton, N., Ogih, M., Okunou, G.,  
Nuyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okunou, G.,  
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,  
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,  
Rivers, M., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,  
Sodergren, E., Sonaik, T., Sparks, A., Stanley, H., Stone, H.,  
Sutton, A., Swatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,  
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,  
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinsón, R., Wang, Q.,  
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,  
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,  
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Welnstock, G., and Gibbs, R.

Direct Submission  
Unpublished  
2 (bases 1 to 180336)  
Worley, K.C.

Submitted (03-FEB-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 180336)  
Worley, K.C.

Direct Submission  
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 12, 2002 this sequence version replaced gi:18846553.

Genome Center  
Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.bhsc.bcm.tmc.edu/>  
Contact: bhsc-help@bcm.tmc.edu

Project Information  
Center project name: GPH

Center clone name: CH230-28717

Summary Statistics  
Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads  
Assembly program: Phrap, version 0.990329

Consensus quality: 138122 bases at least Q40  
Consensus quality: 139518 bases at least Q30

Consensus quality: 144903 bases at least Q20

NOTE: Estimated insert size may differ from sequence length  
(see [http://www.bhsc.bcm.tmc.edu/docs/genbank/draft\\_data.html](http://www.bhsc.bcm.tmc.edu/docs/genbank/draft_data.html)).  
NOTE: This is a 'working draft' sequence. It currently  
consists of 62 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 1178: contig of 1178 bp in length  
\* 1179 1278: gap of unknown length  
\* 1279 2490: contig of 1212 bp in length  
\* 2491 2590: gap of unknown length  
\* 2591 3722: contig of 1132 bp in length  
\* 3723 3822: gap of unknown length

3823 5011: contig of 1189 bp in length  
\* 5012 5111: gap of unknown length  
\* 5112 6446: contig of 1335 bp in length  
\* 6447 7769: gap of unknown length  
\* 7770 7870: contig of 1223 bp in length  
\* 7871 8905: gap of unknown length  
\* 8906 9005: contig of 1036 bp in length  
\* 9006 10193: gap of unknown length  
\* 10194 10293: contig of 1188 bp in length  
\* 10294 11596: gap of unknown length  
\* 11597 11696: contig of 1303 bp in length  
\* 11697 13543: gap of unknown length  
\* 13544 13643: contig of 1847 bp in length  
\* 13644 15725: gap of unknown length  
\* 15726 15826: contig of 2082 bp in length  
\* 15827 17704: gap of unknown length  
\* 17705 17805: contig of 1879 bp in length  
\* 17806 19637: gap of unknown length  
\* 19638 19737: contig of 1833 bp in length  
\* 19738 21484: gap of unknown length  
\* 21485 21585: contig of 1747 bp in length  
\* 21586 23839: gap of unknown length  
\* 23840 23939: contig of 2255 bp in length  
\* 23940 24970: gap of unknown length  
\* 24971 25070: contig of 1030 bp in length  
\* 25071 26476: gap of unknown length  
\* 26477 26577: contig of 1407 bp in length  
\* 26578 29204: gap of unknown length  
\* 29205 29305: contig of 2628 bp in length  
\* 29306 30859: gap of unknown length  
\* 30860 30959: contig of 1555 bp in length  
\* 30960 32587: gap of unknown length  
\* 32588 34480: contig of 1627 bp in length  
\* 34481 34580: gap of unknown length  
\* 34581 35937: contig of 1794 bp in length  
\* 35938 36037: gap of unknown length  
\* 36038 37577: contig of 1356 bp in length  
\* 37578 37677: gap of unknown length  
\* 37678 38964: contig of 1540 bp in length  
\* 38965 39064: gap of unknown length  
\* 39065 41372: contig of 1288 bp in length  
\* 41373 41471: gap of unknown length  
\* 41472 43481: contig of 2307 bp in length  
\* 43482 43581: gap of unknown length  
\* 43582 45439: contig of 2010 bp in length  
\* 45440 45539: gap of unknown length  
\* 45540 48070: contig of 1858 bp in length  
\* 48071 48170: gap of unknown length  
\* 48171 48659: contig of 2530 bp in length  
\* 48660 49876: gap of unknown length  
\* 49877 49976: contig of 1707 bp in length  
\* 49977 52658: gap of unknown length  
\* 52659 52759: contig of 2682 bp in length  
\* 52760 54635: gap of unknown length  
\* 54636 54735: contig of 1877 bp in length  
\* 54736 57623: gap of unknown length  
\* 57624 57722: contig of 2887 bp in length  
\* 57723 59661: gap of unknown length  
\* 59662 61953: contig of 1939 bp in length  
\* 61954 62053: gap of unknown length  
\* 62054 64139: contig of 2192 bp in length  
\* 64140 64239: gap of unknown length  
\* 64240 67756: contig of 2086 bp in length  
\* 67757 67856: gap of unknown length  
\* 67857 70906: contig of 3517 bp in length  
\* 70907 71006: gap of unknown length  
\* 71007 73236: contig of 3050 bp in length  
\* 73237 73337: gap of unknown length  
\* 73338 75973: contig of 2230 bp in length  
\* 75974 76074: gap of unknown length  
\* 76075 78833: contig of 2637 bp in length  
\* 78834 78833: contig of 2760 bp in length

```

* 78834 78933: gap of unknown length
* 78934 81146: contig of 2213 bp in length
* 81147 81246: gap of unknown length
* 81247 84314: contig of 3268 bp in length
* 84315 84614: gap of unknown length
* 84615 88398: contig of 3784 bp in length
* 88399 88498: gap of unknown length
* 88499 91447: contig of 2949 bp in length
* 91448 91547: gap of unknown length
* 91548 95022: contig of 3475 bp in length
* 95023 95122: gap of unknown length
* 95123 99168: contig of 4046 bp in length
* 99169 99268: gap of unknown length
* 99269 102681: contig of 3413 bp in length
* 102682 102781: gap of unknown length
* 102782 106573: contig of 3791 bp in length
* 106573 112143: gap of unknown length
* 112143 112143: contig of 5471 bp in length
* 112144 112243: gap of unknown length
* 112244 116064: contig of 3821 bp in length
* 116065 116164: gap of unknown length
* 116165 120276: contig of 4112 bp in length
* 120277 120376: gap of unknown length
* 120377 123959: contig of 3583 bp in length
* 123960 124059: gap of unknown length
* 124060 128058: contig of 3999 bp in length
* 128059 128158: gap of unknown length
* 128159 132524: contig of 4366 bp in length

```

## Alignment Scores:

```

Pred. No.: 0.252 Length: 180336
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.37% Indels: 0
DB: 2 Gaps: 0

```

US-09-854-133-586 (1-97) x AC108976 (1-180336)

QY 23 Arglystysgltarglystlystargltunrlyst 34

DB 63702 AGAAGAGAAAGAGAGAGAGAGAGAGAGAGAG 63667

## RESULT 44

AC026367/c

## LOCUS

Homo sapiens 12 BAC RP11-131L12 (Roswell Park Cancer Institute

Human BAC Library) complete sequence.

## ACCESSION

AC026367.14 GI:14669931

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 181445)  
 Muzny,D.M., Adams,C., Adio-Obiola,B., Ali-Osman,F.R., Allen,C.,  
 Alshrocks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbara,J.,  
 Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J.,  
 Bowtle,S., Briteva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C.,  
 Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,  
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,  
 Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C., Cleveland,C.D.,  
 Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,  
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,  
 Demn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,C.,  
 Dugan-Rocha,S., Durbin,K.U., Earhart,C., Edgar,D., Edwards,C.C.,  
 Elhaj,C., Emerling,S., Escotto,M., Falls,T., Ferriguto,D.,  
 Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,D.,  
 Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W.,  
 Gunaratne,P., Hale,S., Hamilton,K., Han,J., Harris,C., Harris,K.,  
 Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O.,  
 Hodgson,A., Hognes,M., Holloway,C., Hollins,B., Homisi,F.,

## REFERENCE

## AUTHORS

## JOURNAL

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

## TITLE

## JOURNAL

## AUTHORS

Howard,S., Huber,J., Hulyk,S., Hume,J., Ioshikhes,I., Jackson,L.E.,  
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,  
 Karlsson,E., Kelly,S., Khan,U., King,L., Korah,J., Kovar,C.,  
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lee,E., Lewis,J.C.,  
 Lewis,L., Li,J., Li,Z., Licharge,O., Lieu,C., Liu,D., Liu,W.,  
 Louised,H., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,  
 Ma,J., Maheshwari,M., Mapa,P., Marondel,I., Martin,R.,  
 Martindale,A., Martinez,E., Massey,E., Maubiney,E., McLeod,M.P.,  
 Meador,M., Mel,G., Merscher,S., Metzger,M., Miller,A., Miner,G.,  
 Miner,Z., Mitchell,T., Monabaci,K., Montgomery,K.T., Morgan,N.,  
 Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,M.,  
 Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokkwo,S.,  
 Ogulu,M., Okunoye,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,  
 Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,  
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojibokan,I., Rolfe,M.,  
 Ruiz,S., Saverly,G., Scherer,S., Scott,G., Shen,H., Shim,C.,  
 Shooshari,N., Slisson,I., Sodergren,E., Sonalke,T., Sparks,A.,  
 Stanley,H., Stone,H., Sutton,A., Syatek,A., Tabor,P., Tamerisa,A.,  
 Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B.,  
 Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D.,  
 Vinson,R., Wall,R., Wang,S., Ward-Moore,S., Warren,R.,  
 Washington,C., Wallington,S., Williams,G., Williamson,A.,  
 Wleczky,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J.,  
 Zorilla,S., Zuchelapatti,R. and Gdbbs,R.  
 Direct Submission  
 Unpublished  
 2 (bases 1 to 181445)  
 Worley,K.C.  
 Direct Submission  
 Submitted (22-MAR-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 181445)  
 Worley,K.C.  
 Direct Submission  
 Submitted (11-JUL-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Jul 11, 2001 this sequence version replaced g1:13493004.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
 gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 features listing.

ANNOTATION OF FEATURES:  
 STSs are identified using ePCR (genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.  
 Repeats are identified using RepeatMasker (A. Smit and P. Green,  
 unpublished.) for Human and Mouse sequences.  
 Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintain sequence  
 continuity across the splice junctions. Sequences that are not  
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum  
 standard of double strand coverage with a minimum of 2 clones and 2  
 reads with no ambiguities or 2 chemistries with a minimum of 2  
 clones and 3 reads with no ambiguities. If the sequence quality for  
 a region does not meet this standard, it will be indicated in the  
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
 standards - estimated error rate less than 1 per 10,000 bases.  
 Reports of lowest quality individual bases and measures of base  
 quality are listed below. Description of the metrics can be found

QUALSTAT-REPORT-----

```

Fullap values in estimate:      180115
Average error rate (RCM-Phrap estimate): 0.000170582
Fraction of Phrap values less than 40 : 0.0184993
Number of consensus changing edits: 0
Number of N's in consensus : 0
0

```

Position	Original+Context	Consensus	changing	edits
1715	agctcgtctc (n) gcttcctcctga	agctcgtctc (c) gactcctcctga		agctcgtctc (c) gactcctcctga
1743	tgatcacccc (n) ccttcgtcctc	tgatcacccc (a) ccttcgtcctc		tgatcacccc (a) ccttcgtcctc
1806	ggagtgatga (n) taatttttaa	ggagtgatga (t) taatttttaa		ggagtgatga (t) taatttttaa
1843	gtaactgtccc (n) atgtatggaaa	gtaactgtccc (a) atgtatggaaa		gtaactgtccc (a) atgtatggaaa
1852	cnatgtatgaa (n) nacttctggaaa	cnatgtatgaa (a) nacttctggaaa		cnatgtatgaa (a) nacttctggaaa
1866	tgggaaataac (n) catattatga	tgggaaataac (a) catattatga		tgggaaataac (a) catattatga
1909	ttggtctcaaaa (n) nmdtatcccg	ttggtctcaaaa (c) nmdtatcccg		ttggtctcaaaa (c) nmdtatcccg
1911	ggctcaaaan (n) ngttaatcccg	ggctcaaaan (c) ngttaatcccg		ggctcaaaan (c) ngttaatcccg
2049	gtctcaaaan (n) gtaatcccg	gtctcaaaan (c) gtaatcccg		gtctcaaaan (c) gtaatcccg
2067	gtctcgtgtaa (n) cccagatctact	gtctcgtgtaa (c) cccagatctact		gtctcgtgtaa (c) cccagatctact
2114	actccaggaga (a) ttggagcagga	actccaggaga (c) ttggagcagga		actccaggaga (c) ttggagcagga
2122	gttcagatga (n) ccagagatngc	gttcagatga (c) ccagagatngc		gttcagatga (c) ccagagatngc
2125	ganccgagaat (n) gmcacatcgc	ganccgagaat (c) gmcacatcgc		ganccgagaat (c) gmcacatcgc
2133	cccgagatagc (n) ccaatcgtnc	cccgagatagc (c) ccaatcgtnc		cccgagatagc (c) ccaatcgtnc
2137	gmcacatcgc (n) cctccagctctg	gmcacatcgc (c) cctccagctctg		gmcacatcgc (c) cctccagctctg
2163	gtgggttgcct (n) taagatllta	gtgggttgcct (c) taagatllta		gtgggttgcct (c) taagatllta
4307	atgtgtctctc (n) taagaatttc	atgtgtctctc (a) taagaatttc		atgtgtctctc (a) taagaatttc
36661	ggaagcggaan (n) tncagtgatag	ggaagcggaan (c) tncagtgatag		ggaagcggaan (c) tncagtgatag
36662	gagcgcggaan (n) tncagtgatag	gagcgcggaan (c) tncagtgatag		gagcgcggaan (c) tncagtgatag
36664	ggcggaant (n) ccaatgagctg	ggcggaant (c) ccaatgagctg		ggcggaant (c) ccaatgagctg
51326	gtgggttgcct (n) gcttcgtatc	gtgggttgcct (c) gcttcgtatc		gtgggttgcct (c) gcttcgtatc
51334	atngctttag (n) tcaagtgatc	atngctttag (c) tcaagtgatc		atngctttag (c) tcaagtgatc
51584	atgggagaaat (n) tttttttttt	atgggagaaat (c) tttttttttt		atgggagaaat (c) tttttttttt
52661	agagttgagct (n) gatttttata	agagttgagct (c) gatttttata		agagttgagct (c) gatttttata
52675	ttctagacct (n) gtagctctgc	ttctagacct (c) gtagctctgc		ttctagacct (c) gtagctctgc
57080	caacaatcca (n) ctaacatcgaa	caacaatcca (c) ctaacatcgaa		caacaatcca (c) ctaacatcgaa
64017	aatataaaaa (n) gggacaatga	aatataaaaa (c) gggacaatga		aatataaaaa (c) gggacaatga
86177	tttttttttt (n) gggcagatct	tttttttttt (c) gggcagatct		tttttttttt (c) gggcagatct
92345	catltaatt (n) aaccatcatc	catltaatt (c) aaccatcatc		catltaatt (c) aaccatcatc
97328	aaaaaaaaaa (n) aaaaagaag	aaaaaaaaaa (a) aaaaagaag		aaaaaaaaaa (a) aaaaagaag
102654	gctaattgtt (n) taatttttag	gctaattgtt (c) taatttttag		gctaattgtt (c) taatttttag
102701	agctcgtctc (n) gnatctctga	agctcgtctc (c) gnatctctga		agctcgtctc (c) gnatctctga
102718	gctcgtctctg (n) atctctgact	gctcgtctctg (c) atctctgact		gctcgtctctg (c) atctctgact
102771	ctgactctcag (n) ngatcacacc	ctgactctcag (c) ngatcacacc		ctgactctcag (c) ngatcacacc
102774	tgactctcag (n) gattatcccg	tgactctcag (c) gattatcccg		tgactctcag (c) gattatcccg
102784	gcccgagctc (n) nnaattatcc	gcccgagctc (c) nnaattatcc		gcccgagctc (c) nnaattatcc
102786	cccgagctcgt (n) naattttttt	cccgagctcgt (c) naattttttt		cccgagctcgt (c) naattttttt
102820	cggcgccagc (n) gattcttttt	cggcgccagc (c) gattcttttt		cggcgccagc (c) gattcttttt
113281	cggcgccagc (n) gattcttttt	cggcgccagc (c) gattcttttt		cggcgccagc (c) gattcttttt
113282	agatggaggtc (n) naanaacmg	agatggaggtc (c) naanaacmg		agatggaggtc (c) naanaacmg
113284	gatggaggtc (n) anaacmgng	gatggaggtc (c) anaacmgng		gatggaggtc (c) anaacmgng
113285	ggaggtttmaa (n) nacmgngcc	ggaggtttmaa (c) nacmgngcc		ggaggtttmaa (c) nacmgngcc
113299	ggaggtttmaa (n) acmgngcca	ggaggtttmaa (c) acmgngcca		ggaggtttmaa (c) acmgngcca
113300	ttannaaag (n) hgagcagagct	ttannaaag (c) hgagcagagct		ttannaaag (c) hgagcagagct
155971	ttannaaag (n) gggcagagctc	ttannaaag (c) gggcagagctc		ttannaaag (c) gggcagagctc
	ttggagacat (n) tcttagcttc	ttggagacat (c) tcttagcttc		ttggagacat (c) tcttagcttc

Distribution of Quality &lt; 40 Bases

# bases
1000
900
800
700
600
500

	Phrap Value	Range	35	40
400	*		*	*
300	*		*	*
200	*		*	*
100	*		*	*
0	*		*	*
	5	10	15	20
			25	30

## FEATURES

...

3

ਪੰਨਾ ੧

rep

Prep No

...more  
...more

## Best Local

**DB:**

HS-09-854

Q

AC118757

DEFINITION

## VERSION

## SOURCE

**QUESTIONS**

100

## AUTHORS

JOURNAL

**ЗАПЫСЬ**

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talmas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

**TITLE**  
JOURNAL  
REFERENCE  
AUTHORS

Query Match: 12.37% Indels: 0  
DB: 2 Gaps: 0  
US-09-854-133-586 (1-97) x AC118757 (1-183011)  
QY 23 ArgLySLysGLUArgLySLysSLysSLysArgGLUArgLys 34  
DB 7219 AGAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 7184

Search completed: May 11, 2003, 17:19:47  
Job time : 2680.93 secs

**TITLE**  
JOURNAL  
COMMENT

Submitted (23-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Aug 23, 2002 this sequence version replaced g1:22267634.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WtBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L26244

Center clone name: 905\_K\_4

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 116580: contig of 116580 bp in length  
\* 116581 116680: gap of 100 bp  
\* 116681 183011: contig of 66331 bp in length.

Location/Qualifiers

1. 183011

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/chromosome="18"

/map="18"

/clone="RP11-905K4"

/clone\_lib="RC1-11 Human Male BAC"

BASE COUNT 49244 a 37615 c 38929 g 56894 t 329 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.255 Length: 183011  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:14:30 ; Search time 207.735 Seconds  
(without alignments)

1051.553 Million cell updates/sec

Title: US-09-854-133-586

Sequence: 1 EVEVSRDHSALSDSETLSQT.....LFGCLPMATRSHLGRKCS 97

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4368722

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 60 summaries

Command line parameters:

-MODEL=frame+2.p2n.model -DEV=xlh  
-O/cg2.1/USPTO.spool/US09854133/unat\_05052003.174132\_694/app\_query.fasta.1.462  
-Db=N.Geneseq.101002 -QFWT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=60 -DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1 -ALIGN=45 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEP=SIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09854133.ecgn.1.1\_275\_etunat\_05052003.174132\_694 -NCPU=6 -ICPU=3  
-NO.XLPHY -NO.MMAP -LARGESOURCE -NEG.SCORES=0 -WAIT -LONLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

N.Geneseq.101002: \*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT: \*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	97	100.0	337	23	Human lung tumour-
2	96	99.0	5981	23	Human lung tumour-
3	93	95.9	2239	23	Human lung tumour-
4	37	38.1	575	24	Human colon tumour
5	37	38.1	596	24	Human colon cancer
6	37	38.1	800	20	Human gene express
7	37	38.1	1958	22	Human EST-derived
8	37	38.1	1958	22	Human immune/haema
9	36	37.1	123	22	Human musculoskele
10	36	37.1	29607	22	Human cardiovascular
11	36	37.1	32190	22	Human immune/haema
12	36	37.1	32190	22	Human immune/haema
13	36	37.1	32190	22	Human secreted pro
14	36	37.1	32190	22	Human secreted exp
15	36	37.1	32190	22	Human secreted exp
16	36	37.1	32190	22	Human secreted exp
17	36	37.1	32190	22	Human secreted exp
18	36	37.1	32190	22	Human secreted exp
19	36	37.1	32190	22	Human secreted exp
20	36	37.1	32190	22	Human secreted exp
21	36	37.1	32190	22	Human secreted exp
22	36	37.1	32190	22	Human secreted exp
23	36	37.1	32190	22	Human secreted exp
24	36	37.1	32190	22	Human secreted exp
25	36	37.1	32190	22	Human secreted exp
26	36	37.1	32190	22	Human secreted exp
27	36	37.1	32190	22	Human secreted exp
28	36	37.1	32190	22	Human secreted exp
29	36	37.1	32190	22	Human secreted exp
30	36	37.1	32190	22	Human secreted exp
31	36	37.1	32190	22	Human secreted exp
32	36	37.1	32190	22	Human secreted exp
33	36	37.1	32190	22	Human secreted exp
34	36	37.1	32190	22	Human secreted exp
35	36	37.1	32190	22	Human secreted exp
36	36	37.1	32190	22	Human secreted exp
37	36	37.1	32190	22	Human secreted exp
38	36	37.1	32190	22	Human secreted exp
39	36	37.1	32190	22	Human secreted exp
40	36	37.1	32190	22	Human secreted exp
41	36	37.1	32190	22	Human secreted exp
42	36	37.1	32190	22	Human secreted exp
43	36	37.1	32190	22	Human secreted exp
44	36	37.1	32190	22	Human secreted exp
45	36	37.1	32190	22	Human secreted exp
46	36	37.1	32190	22	Human secreted exp
47	36	37.1	32190	22	Human secreted exp
48	36	37.1	32190	22	Human secreted exp
49	36	37.1	32190	22	Human secreted exp
50	36	37.1	32190	22	Human secreted exp
51	36	37.1	32190	22	Human secreted exp
52	36	37.1	32190	22	Human secreted exp
53	36	37.1	32190	22	Human secreted exp
54	36	37.1	32190	22	Human secreted exp
55	36	37.1	32190	22	Human secreted exp
56	36	37.1	32190	22	Human secreted exp
57	36	37.1	32190	22	Human secreted exp
58	36	37.1	32190	22	Human secreted exp
59	36	37.1	32190	22	Human secreted exp
60	36	37.1	32190	22	Human secreted exp

## ALIGNMENTS

RESULT 1  
AAD23462 standard; CDNA: 337 BP.  
ID AAD23462

```

XX AD23462;
AC
XX 26-FEB-2002 (first entry)
DE Human lung tumour-specific 20E10 5' cDNA.
XX Human lung tumour protein; immunostimulant; cytostatic; gene therapy;
XX antisense-therapy; vaccine; immune response; lung cancer; 20E10; ss.
OS Homo sapiens.
XX
XX WO200172295-A2.
XX
XX 04-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US09991.
XX
XX 29-MAR-2000; 2000US-0538037.
XX 05-JUN-2000; 2000US-0588937.
XX 18-AUG-2000; 2000US-0640878.
XX 22-SEP-2000; 2000US-234517P.
XX 01-NOV-2000; 2000US-0704512.
XX 14-DEC-2000; 2000US-0738973.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY,
XX Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX WPI; 2001-639201/73.
XX
XX New human lung-specific polynucleotides and polypeptides for the
XX diagnosis and treatment of disease e.g. lung cancer -
XX
XX Claim 1; Page 334; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and
XX their corresponding cDNA molecules. Lung tumour-specific proteins and
XX their antigen-presenting cells are useful for stimulating and/or
XX the development of cancer. The invention also relates to a composition
XX useful for stimulating an immune response, and for inhibiting
XX lung tumour specific oligonucleotide is useful in gene therapy and for
XX diagnosis, detection and treatment of lung cancer. The present sequence
XX is a cDNA encoding human lung tumour-specific protein.
XX
XX Sequence 337 BP; 103 A; 60 C; 93 G; 81 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 1 22e-89 Length: 337
XX Score: 97.00 Matches: 97
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 23 Gaps: 0
XX
XX US-09-854-133-586 (1-97) x AAD23462 (1-337)
XX
XX 1 GLUVALGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr
XX 5 GAGGTTGAAGTGAAGCAGAGATCATGCCACCTGGGTGACAGTGAAGCTGTCTCAACA
XX 21 GluLeuArgGlyGlyGluArgGlyGlyGlyGluArgGlyGlyGlyGlyGlyGlyGly
XX 65 GAATTAAGGAAAAAGAAAAAGAAAAAGAGAGAGAGAAATTCACAGCAATTTGTGGC
XX 41 TLeaSphelellelephetrillephetrilleleuLeuPheSerHisIstTPIlegin
XX 125 ATGAGATTTCATATCTGCAATTTTGTGATTTCTTTGTTCTTCATCAGTTCAG
XX 61 GluSerleuLeuLeuProSerProLeuGluValThrGysArgGluMetLeuThrGly

```

```

DB 185 GAACCTGTGTGTGTCACCATCTCCAAAGAGGTACCTGCAGGAAATGTTAACGGCA
XX
XX 81 GlycLeuProTyrPalatThrArgSerHisLeuGlyArgArgGlyCysSer 97
XX
XX 245 GGTGCTCTTCCCTGGCAACAAGAGCCACCTGGGACAGAGAAAGTGCAGC 295
XX
XX RESULT 2
XX AAD23461
XX ID AAD23461 standard; cDNA; 5981 BP.
XX
XX AC AAD23461;
XX
XX 26-FEB-2002 (first entry)
XX
XX Human lung tumour-specific 14F10 full length cDNA.
XX
XX Human lung tumour protein; immunostimulant; cytostatic; gene therapy;
XX antisense-therapy; vaccine; immune response; lung cancer; 14F10; ss.
XX
XX Homo sapiens.
XX
XX WO200172295-A2.
XX
XX 04-OCT-2001.
XX
XX 28-MAR-2001; 2001WO-US09991.
XX
XX 29-MAR-2000; 2000US-0538037.
XX 05-JUN-2000; 2000US-0588937.
XX 18-AUG-2000; 2000US-0640878.
XX 22-SEP-2000; 2000US-234517P.
XX 01-NOV-2000; 2000US-0704512.
XX 14-DEC-2000; 2000US-0738973.
XX
XX (CORI-) CORIXA CORP.
XX
XX Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indirias CY,
XX Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX WPI; 2001-639201/73.
XX
XX New human lung-specific polynucleotides and polypeptides for the
XX diagnosis and treatment of disease e.g. lung cancer -
XX
XX Claim 1; Page 332-334; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and
XX their corresponding cDNA molecules. Lung tumour-specific proteins and
XX their antigen-presenting cells are useful for stimulating and/or
XX the development of cancer. The invention also relates to a composition
XX useful for stimulating an immune response, and for inhibiting
XX lung tumour specific oligonucleotide is useful in gene therapy and for
XX diagnosis, detection and treatment of lung cancer. The present sequence
XX is a cDNA encoding human lung tumour-specific protein.
XX
XX Sequence 5981 BP; 1858 A; 1029 C; 1098 G; 1996 T; 0 other.
XX
XX Alignment Scores:
XX Pred. No.: 1 85e-87 Length: 5981
XX Score: 96.00 Matches: 96
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 98.97% Indels: 0
XX DB: 23 Gaps: 0
XX
XX US-09-854-133-586 (1-97) x AAD23461 (1-5981)
XX
XX 2 ValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr
XX 3 GTTGAAGTGAAGCAGAGATCATGCCACCTGGGTGACAGTGCAGACTGTCTCAACAAGAA
XX 22 LeuArgGlyGlyGluArgGlyGlyGlyGluArgGlyGlyGlyGlyGlyGlyGly

```

```

|||||
DB 63 TTRAGAGAAAAAGAAAGAAAAAGAGAGAGAAATTCAGGCATGTGCATTA 122
QY 42 Aappheillelephetriphelelelelepheserishsttrpilleglu 61
DB 123 GATTATATCATATTCGATTTTGTGATTTCTTTTCATCATCGATTCAGGAA 182
QY 62 SerleuleucysproproserprolysGluValThrCysArgGluMetleuthrGly 81
DB 183 AGCTGTGTGTCCACCATCTCCAAAGAGAGTTACCTGCAGGGAATGTTAACGGGAGGC 242
QY 82 CysleuprottrpalaThrArgSerHisleuGlyArgArglyscysSer 97
DB 243 TGCTCTCCCTGGGCAACAGAGCACCCTGGCAGAGAAAGTGCACG 290

RESULT 3
AAD23460
ID AAD23460 standard; cDNA; 2239 BP.
XX
AC AAD23460;
XX
DT 26-FEB-2002 (first entry)
DE Human lung tumour-specific 19A4 cDNA.
XX
XX Human; lung tumour protein; immunostimulant; cytostatic; gene therapy;
KM antisense-therapy; vaccine; immune response; lung cancer; 19A4; ss.
XX
OS Homo sapiens.
XX
PN WO200172295-A2.
XX
PD 04-OCT-2001.
XX
PF 28-MAR-2001; 2001WO-US09991.
XX
PR 29-MAR-2000; 2000US-0538037.
PR 05-JUN-2000; 2000US-0588937.
PR 18-AUG-2000; 2000US-0640878.
PR 22-SEP-2000; 2000US-234517P.
PR 01-NOV-2000; 2000US-0704512.
PR 14-DEC-2000; 2000US-0738973.
XX
PA (CORI-) CORIXA CORP.
XX
PI Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
PI Henderson RA, Fling SP, Algate PA, Elliot M, Mannion J, Kalos MD;
XX
XX WPI: 2001-639201/73.
XX
DR New human lung-specific polynucleotides and polypeptides for the
PT diagnosis and treatment of disease e.g. lung cancer -
XX
PS Claim 1; Page 332; 378pp; English.
XX
XX The invention relates to isolated lung tumour-specific proteins and
CC their corresponding cDNA molecules. Lung tumour-specific proteins and
CC their antigen-presenting cells are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, and for inhibiting
CC the development of cancer. The invention also relates to a composition
CC useful for stimulating an immune response, and for treating cancer. The
CC lung tumour specific oligonucleotide is useful in gene therapy and for
CC diagnosis, detection and treatment of lung cancer. The present sequence
CC is a cDNA encoding human lung tumour-specific protein.
XX
SQ Sequence 2239 BP; 619 A; 444 C; 493 G; 683 T; 0 other;

Alignment Scores:
Pred. No.: 8, 3e-85 Length: 2239
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.88% Indels: 0

```

```

DB: 23 Gaps: 0
US-09-854-133-586 (1-97) x AAD23460 (1-2239)
QY 1 GluValGluValSerArgAspHisAlaSerleuGlyAspSerGluThrLeuSerGluThr 20
DB 2 GAGGTAAAGTACAGACAGATCATGCGCAGCGTGGTACAGTACAGTCTGTCAACA 61
QY 21 GluLeuArglyslsGluArglyslsLysArgGluArgGlyPheGlnAlaAsnCysGly 40
DB 62 GAATTAAGGAAAAAGAAAGAAAGAAAGAGAGAGGAAATTCAGGCCAATTTGCGC 121
QY 41 Ileasphelleillephetripheleleleleupheserishsttrpilleglu 60
DB 122 ATGATTTTATCATATTCGTGATTTTGTGATTTCTTTGTTTCATCATCTGATTCAG 181
QY 61 GluSerleuLeuCyproproserprolysGluValThrCysArgGluMetleuthrGly 80
DB 182 GAAAGCCTGTTGTGTCACCATCTCCAAAGAGAGTTACCTGCAGGGAATGTTAACGGGGA 241
QY 81 GlyCysleuprottrpalaThrArgSerHisleuGlyArg 93
DB 242 GCCTGCTCTCCCTGGGCAACAGAGCCACCTGGGCAAG 280

RESULT 4
ABL37718
ID ABL37718 standard; cDNA; 575 BP.
XX
AC ABL37718;
XX
DT 08-APR-2002 (first entry)
DE Human colon tumour antigen polynucleotide SEQ ID NO:1307.
XX
XX Human; colon cancer; colon tumour antigen; cytostatic; vaccine;
KM colon tumour metastatic antigen; diagnosis; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200196388-A2.
XX
PD 20-DEC-2001.
XX
PF 08-JUN-2001; 2001WO-US18557.
XX
PR 09-JUN-2000; 2000US-210899P.
PR 20-FEB-2001; 2001US-270216P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Jiang Y, Harlocker SL, Secrist H;
XX
XX WPI: 2002-114514/15.
XX
DR Novel isolated colon tumor polynucleotide differentially expressed in
PT colon tumor or colon metastatic tumor and polypeptides encoded by them,
PT useful for inhibiting development of cancer in patient -
XX
PS Claim 1; SEQ ID 1307; 105pp; English.
XX
XX ABL36412 to ABL38645 represent human colon tumour antigen cDNA clones (1)
CC which were isolated from human colon tumour and colon metastatic tumour
CC cDNA libraries. (1) have cytostatic activity and can be used in vaccine
CC production. (1) can be used for stimulating and/or expanding T cells
CC specific for a tumour protein on contact with the T cells. They are also
CC useful for inhibiting the development of cancer in a patient. (1) can be
CC used as probes or primers for nucleic acid hybridisation, for preparing
CC mutant species primers, or primers for use in genetic constructions. (1)
CC can be used in the diagnosis of a colon tumour.
XX
SQ Sequence 575 BP; 151 A; 122 C; 153 G; 141 T; 8 other;

Alignment Scores:

```

Pred. No.: 1,9e-28  
 Score: 37.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 38.14%  
 DB: 24  
 Matches: 575  
 Conservatives: 37  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-09-854-133-586 (1-97) x ABL37718 (1-575)

OY 61 GluSerLeuLeuCySPProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 DB 256 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGAAATGTTAACGGGA 315  
 OY 81 GlyCysLeuProTPrAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
 DB 316 GGCTGCTTCCCTGGCAACAGAGCCACCTGGGAGGAAAGTGCAGC 366

RESULT 5

ID ABR27651/C  
 ABK27651 standard; CDNA; 596 BP.

AC ABR27651;

DT 09-APR-2002 (first entry)

DE Human colon cancer expressed sequence tag, Seq ID no 88.

KW Human; colon cancer; T cell expansion; tumour; EST; gene; ss;

OS Homo sapiens.

PN W0200196390-A2.

PD 20-DEC-2001.

PF 08-JUN-2001; 2001WO-US18577.

PR 09-JUN-2000; 2000US-210821P.

PR 18-DEC-2000; 2000US-256571P.

PR 10-MAY-2001; 2001US-290240P.

PA (CORI-) CORIXA CORP.

PI Jiang Y, Hepler WT, Clapper JD, Wang A, Secrist H;

DR WPI; 2002-139708/18.

Novel isolated polynucleotide encoding a polypeptide comprising a portion of colon tumour protein, useful for detection, diagnosis and therapy of human colon cancer.

Claim 1; Page 174; 220pp; English.

The invention relates to an isolated polynucleotide (I) encoding a polypeptide (II) comprising at least a portion of a colon tumour protein. (I), (II) and antibody (III) to (II) are useful for determining the presence of a cancer in a patient. (I), (II) or antigen presenting cells expressing (I) is useful for stimulating and/or expanding T cells specific for a tumour protein, by contacting T cells with (I), (II) or antigen-presenting cells that express (I), under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. (I), (II), or antigen presenting cells that express (II) are useful for treating colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated from a patient with (I), (II) or antigen presenting cells that express (II), such that T cells proliferate, and administering to the patient an effective amount of the proliferated T cells, thus inhibiting the development of a cancer in the patient. (I) or (II) is useful in vaccines and pharmaceutical compositions for prevention and treatment of colon malignancies and for the diagnosis and monitoring of such cancers. (I), (II) or (III) is useful for detection, diagnosis and/or therapy of human colon cancer. (I) is useful as a probe or primer for nucleic acid hybridisation, and in the design and preparation of

CC ribozyme molecules for inhibiting expression of (II) in tumour cells.  
 CC ABR27564-ABK27807 represent novel human colon cancer coding  
 CC sequences and primers of the invention.  
 XX

SQ Sequence 596 BP; 160 A; 151 C; 126 G; 158 T; 1 other.

Alignment Scores:

Pred. No.: 1,97e-28  
 Score: 37.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 38.14%  
 DB: 24  
 Matches: 596  
 Conservatives: 37  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-09-854-133-586 (1-97) x ABR27651 (1-596)

OY 61 GluSerLeuLeuCySPProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 DB 367 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACCTGCAGGAAATGTTAACGGGA 308

OY 81 GlyCysLeuProTPrAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
 DB 307 GGCTGCTTCCCTGGCAACAGAGCCACCTGGGAGGAAAGTGCAGC 257

RESULT 6

ID AAZ16609  
 AAZ16609 standard; CDNA; 800 BP.

AC AAZ16609;

DT 12-OCT-1999 (first entry)

DE Human gene expression product CDNA sequence SEQ ID NO:4079.

KW Human; gene; gene expression product; diagnosis; therapy; probe;

KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

OS Homo sapiens.

PN W09938972-A2.

PD 05-APR-1999.

PF 28-JAN-1999; 99WO-US01619.

PR 03-APR-1998; 98US-0080666.

PR 28-JAN-1998; 98US-0072910.

PR 31-MAR-1998; 98US-0075954.

PR 03-APR-1998; 98US-0080114.

PA (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Crvenjakov R, Dickson M, Drmanac R, Drmanac S;  
 PI Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;  
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
 PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;

DR WPI; 1999-494092/41.

Novel human genes and their expression products which are differentially expressed in different cell types

Claim 1; Page 1934; 247pp; English.

The present invention describes a library of human polynucleotides comprising the sequences given in AAZ12532 to AAZ17779. Also described is a method of detecting differentially expressed genes correlated with the cancerous state of a mammalian cell, comprising detecting at least one differentially expressed gene product in a test sample from a cell

CC suspected of being cancerous, where the gene product is encoded by one  
 CC of the 5248 polynucleotide sequences given in AA12352 to AA21779. The  
 CC polynucleotides can be used as a source of primers and probes, which can  
 CC be used for a variety of purposes, e.g. detection of expression levels,  
 CC mapping, tissue typing or profiling, forensics, genetic analysis and  
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
 CC can be used for raising antibodies for experimental, diagnostic and  
 CC therapeutic purposes. The polynucleotides may also be used to construct  
 CC arrays for diagnostics (which may be used to determine function of an  
 CC encoded protein); and to detect differences in expression levels between  
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
 CC identify a genetic predisposition or susceptibility to a disease such as  
 CC cancer). The polynucleotides of the invention are especially used in the  
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
 CC and lung cancer. The polynucleotides can also be used to screen for  
 CC peptide analogues and antagonists.

XX  
 SQ Sequence 800 BP; 199 A; 166 C; 202 G; 211 T; 22 other;

# Alignment Scores:

Pred. No.:	2,59e-28	Length:	800
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	20	Gaps:	0

US-09-854-133-586 (1-97) x AA216609 (1-800)

QY 61 GlusertleuLeuCyProProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 |||||||

DB 344 GAAAGCCTGTGTGTCCACATCTCCAAAGAGGTTACTGAGGAAATGTTAAACGGGA 403  
 |||||||

QY 81 GlyCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
 |||||||

DB 404 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGAGGAAAGTGCAGC 454  
 |||||||

# RESULT 7

ID ABA09201 standard; cDNA; 1958 BP.

XX ABA09201;

DT 11-JAN-2002 (first entry)

DE Human cysteine/Glu transporter homologue cDNA, SEQ ID NO:977.

XX Human; cytokine; cell proliferation; cell differentiation; growth factor;  
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;  
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;  
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;  
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;  
 KW chronic inflammatory condition; proliferative retinopathy;  
 KW atherosclerosis; coronary heart disease; arterial ischemia;  
 KW bone disorder; osteoporosis; vascular growth disorder;  
 KW tissue regeneration; wound healing; infection; immune disorder;  
 KW cell culture; drug screening; gene therapy; antiinflammatory;  
 KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;  
 KW cystostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;  
 KW antifungal; vulnerary; antilicer; ss.

OS Homo sapiens.

PN WO200157188-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US03800.

PR 03-FEB-2000; 2000US-0496914.  
 PR 27-APR-2000; 2000US-0560875.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

PI WPI; 2001-457740/49.

DR P-PSDB; ABB11957.

PT Human proteins and DNA encoding sequences useful for preventing,  
 PT treating or ameliorating a medical condition in a mammalian subject  
 PT e.g. arthritis and cancer -

PS Claim 1; Page 835; 1963pp; English.

XX Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and  
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The  
 CC invention also relates to vectors and recombinant host cells comprising a  
 CC nucleotide of the invention, methods of producing the novel polypeptides,  
 CC antibodies against the polypeptides, methods of detecting the nucleotides  
 CC or polypeptides in a sample, and methods of identifying compounds which  
 CC bind to polypeptides of the invention. Although novel, many of the  
 CC polypeptides of the invention have homology to known proteins, thereby  
 CC giving an insight into their probable biological activities, and hence  
 CC potential therapeutic applications. The polypeptides of the invention may  
 CC have various activities, including cytokine, cell proliferation or cell  
 CC differentiation activities; stem cell growth factor activity;  
 CC haematopoiesis regulatory activity; tissue growth activity;  
 CC immunomodulatory activity; activin- or inhibin-related activity;  
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or  
 CC thrombolytic activities; receptor or ligand activities; or may be  
 CC involved in oncogenesis, cancer cell proliferation or metastasis.  
 CC Depending on their biological activities, polypeptides and nucleotides of  
 CC the invention are useful for preventing, treating or ameliorating medical  
 CC conditions, e.g., by protein or gene therapy. Such conditions include  
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell  
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),  
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,  
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal  
 CC vascular growth. Polypeptides involved with tissue regeneration and  
 CC repair (or nucleic acids encoding them) may be used to promote wound  
 CC healing (e.g., of burns, incisions and ulcers), while those with  
 CC immunomodulatory activities may be used in the treatment of viral,  
 CC bacterial and fungal infections in addition to immune disorders.  
 CC Polypeptides with growth factor activity may be used in cell cultures to  
 CC promote cell growth. For example, such polypeptides may be used to  
 CC manipulate stem cells in culture to give rise to neuroepithelial cells  
 CC that can be used to augment or replace cells damaged by illness,  
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides  
 CC may also be used in the diagnosis of the above conditions, and in drug  
 CC screening techniques. The present sequence represents a cDNA encoding a  
 CC novel human polypeptide of the invention.

XX SQ Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;

# Alignment Scores:

Pred. No.:	5.97e-28	Length:	1958
Score:	37.00	Matches:	37
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	38.14%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x ABA09201 (1-1958)

QY 61 GlusertleuLeuCyProProSerProLySGluValThrCysArgGluMetLeuThrGly 80  
 |||||||

DB 70 GAAAGCCTGTGTGTCCACATCTCCAAAGAGGTTACTGAGGAAATGTTAAACGGGA 129  
 |||||||

QY 81 GlyCysLeuProTTPalaThrArgSerHisLeuGlyArgArgLysCysSer 97  
 |||||||

DB 130 GGCTGCTTCCCTGGGCAACAGGAGCCACCTGGGAGAGGAAAGTGCAGC 180  
 |||||||

# RESULT 8

AAH98573 standard; cDNA; 1958 BP.

XX AAH98573;  
AC  
XX  
DT 12-OCT-2001 (first entry)  
DE  
XX  
DE Human EST-derived coding sequence SEQ ID NO: 430.  
XX  
XX Human; sheep; pig; cow; fruit fly; yeast; hamster; macaque; horse;  
KW tomato; monkey; dog; sea urchin; expressed sequence tag; EST;  
KW diagnostics; forensic test; gene mapping; genetic disorder;  
KW biodiversity; gene therapy; nutrition; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200154477-A2.  
XX  
PD 02-AUG-2001.  
XX  
PF 25-JAN-2001; 2001WO-US02687.  
XX  
PR 25-JAN-2000; 2000US-0491404.  
PR 17-JUL-2000; 2000US-0617746.  
PR 03-AUG-2000; 2000US-0631451.  
PR 15-SEP-2000; 2000US-0653870.  
XX  
PA (HXSE-) HXSEQ INC.  
XX  
PI Tang YF, Liu C, Zhou P, Qian XB, Wang Z, Chen R, Asundi V;  
PI Cao Y, Dermanac RA, Zhang J, Wehrman T;  
XX  
DR WPT: 2001-476164/51.  
DR P-PSDB: AAM23914.  
XX  
PT Isolated polypeptide for treatment of diseases, diagnostics, raising  
XX antibodies and research use -  
XX  
PS Claim 1; Page 484; 1275pp; English.  
XX  
XX The present invention provides the protein and coding sequences of novel  
CC proteins from a variety of organisms, including human, dog, cat, horse,  
CC cow, pig, hamster, monkey, macaque, yeast, bacteria, fruit fly, sea  
CC urchin and tomato. These were derived from expressed sequence tags (ESTs)  
CC from the organism of interest. They can be used in diagnostics,  
CC forensics, gene mapping, identification of mutations, to assess  
CC biodiversity and for nutritional purposes. The present sequence is a cDNA  
CC of the invention.  
XX  
SQ Sequence 1958 BP; 499 A; 418 C; 424 G; 617 T; 0 other;  
XX  
XX Alignment Scores:  
Pred. No.: 5.97e-28 Length: 1958  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
DB: 22 Gaps: 0  
XX  
US-09-854-133-586 (1-97) x AAH98573 (1-1958)  
QY 61 GluSerLeuLeuGlyProPheProGluValThrGlyArgGluMetLeuThrIle 80  
DB 70 GAAAGCGCTGTGTCACCATCTCCAAAGAGGTTACTGTCAGGGAATTTAAACGGA 129  
QY 81 GlyCysLeuProTAlaThrArgSerHisLeuGlyArgArgIleCysSer 97  
DB 130 GGCTGCTTCCTCCGCGCAACAGAGCACCCTGGCGAGGAAGTGCAGC 180  
RESULT 9  
ID AAK68639 standard; DNA; 123 BP.  
XX AAK68639;  
AC  
XX

DT 06-NOV-2001 (first entry)  
XX  
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:23451.  
XX  
DE Human; Immune; haematopoietic; Immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 14-AUG-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0225759.  
PR 22-AUG-2000; 2000US-0226581.  
PR 22-AUG-2000; 2000US-0226588.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.

```

PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235634.
PR 27-SEP-2000; 2000US-0235636.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

```

```

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251889.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure: SEQ ID NO 23451; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 123 BP; 32 A; 31 C; 37 G; 23 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 4.7e-28 Length: 123
XX Score: 36.00 Matches: 36
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 37.11% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-854-133-586 (1-97) x AAK68639 (1-123)
XX
XX QY 61 GluSerLeuEncysProPserProLyGluValIhrCysArgGluMetLeuThrGly 80
XX |||||||
XX DB 14 GAAAGCCTGTTGTCACCATCTCCAAAGAGAGGTTACCTGCAAGAAATGTTAAACGGA 73
XX
XX QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgGlyCys 96
XX |||||||
XX DB 74 GGCTGCTTCCTCGGCAACAGAGCCACCTGGCGAGAGAAAGTGC 121
XX
XX RESULT 10
XX AAL37261
XX ID AAL37261 standard; DNA: 29607 BP.
XX
XX AC AAL37261;
XX
XX DT 08-JAN-2002 (first entry)
XX
XX DE Human musculoskeletal system related polynucleotide SEQ ID NO 3626.
XX
XX XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
XX KM vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX KM cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX KM neurological disease; infection; human; secreted protein;
XX KM musculoskeletal system; ds.
XX
XX

```

OS Homo sapiens.  
XX  
PN WO200155367-A1.  
XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01338.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 14-JUL-2000; 2000US-0217496.  
PR 26-JUL-2000; 2000US-0218290.  
PR 14-AUG-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 18-AUG-2000; 2000US-0225779.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 06-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 08-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0232088.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 02-OCT-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 13-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239335.  
PR 20-OCT-2000; 2000US-0239337.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244674.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 01-DEC-2000; 2000US-0250160.  
PR 05-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251038.  
PR 06-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 11-DEC-2000; 2000US-0251990.  
PR 05-JAN-2001; 2001US-0259678.  
PA (HUMA-) HUMAN GENOME SCI INC.



XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-451937/48.  
DR  
XX  
XX Isolated polypeptide for treating, preventing and/or prognosing  
PT disorders related to the musculoskeletal system including  
PT musculoskeletal cancers and also for testing and detection e.g.  
PT diagnosis -  
XX  
XX Example 2; SEQ ID NO 3626; 781bp + Sequence listing; English.  
XX  
XX The invention relates to novel genes (AAL3469-AAL3766) and proteins  
CC (AB03087-AB04109) associated with the musculoskeletal system useful  
CC for preventing, treating or ameliorating medical conditions e.g. by  
CC protein or gene therapy. The genes are isolated from a range of human  
CC tissues disclosed in the specification. The nucleic acids, proteins,  
CC antibodies and (ant)agonists are useful in the diagnosis, treatment  
CC and prevention of: (a) cancer, e.g. breast and ovarian cancer and  
CC other cancers of the adrenal gland, bone, bone marrow, breast,  
CC gastrointestinal tract, liver, lung or urogenital; (b) immune  
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic  
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,  
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis;  
CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;  
CC and (f) infectious diseases such as viral, bacterial, fungal and  
CC parasitic infections.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 29607 BP; 8754 A; 4675 C; 5555 G; 10423 T; 0 other;  
Alignment Scores:  
Pred. No.: 0.191 Length: 29607  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
DB: 22 Gaps: 0  
US-09-854-133-586 (1-97) x AAL37261 (1-29607)  
QY 8 HisAlaSerLeuGlyAspSerGluThrLeuSerGln 19  
|||||  
Db 24558 CACGCTACCTGGTGGTACAGTACCTGTCTCAA 24593  
RESULT 11  
ID AAS36709 standard; DNA; 32190 BP.  
XX  
XX AAS36709;  
AC  
XX  
XX 17-DEC-2001 (first entry)  
DE  
XX  
XX Human cardiovascular system antigen genomic DNA SEQ ID No 2209.  
XX  
XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytoskeletal; cardiac; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-fertility.  
XX  
XX Homo sapiens.  
XX  
XX WO200155321-A2.  
PN

XX  
PD 02-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01340.  
XX  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226888.  
PR 23-AUG-2000; 2000US-0227182.  
PR 30-AUG-2000; 2000US-0227009.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 23-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM,  
XX

Sequence IDs: AA535741.1-AB3536942 represent genomic DNA molecules, which encode the cardiovascular system antigens and their associated polypeptides of the invention. Cardiovascular system antigens and their associated polynucleotides are useful in the diagnosis, treatment and prevention of various types of disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological condition can be determined by detecting the presence or absence of a mutation in a cardiovascular system antigen polynucleotide. The treatable disorders include autoimmune diseases such as rheumatoid arthritis, hyperproliferative disorders such as neoplasms of the breast or liver, cardiovascular disorders such as cardiac arrest, cerebrovascular disorders such as cerebral ischaemia, nervous system disorders such as Alzheimer's disease, infections caused by bacteria, viruses and fungi, ocular disorders such as corneal infection, endocrine disorders such as premature labour and infertility, gastrointestinal disorders such as Crohn's disease, renal disorders such as glomerulonephritis and respiratory disorders such as asthma and pleurisy. The polypeptides can also be used to aid wound healing, to prevent skin aging due to sunburn, to maintain organs before transplantation, to regenerate tissues and in chemotaxis.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences).

Pred. No.:	2.15	Length:	32190
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.34%	Indels:	0
DB:	22	Gaps:	0

QY 9 AlaserleugllyaspserglutThrleusergln 19  
|||||  
Db 30055 GCCAGCTGGGCGACGAGACTCTCTCTCA 30087

ID	AAK82863	standard; DNA; 136 BP.
AT		

DT 07-NOV-2001 (first entry)  
 XY

cytostatic: gene therapy: vaccine: metastasis: de

Homo sapiens.

PN WO200157182-A2.

09-AUG-2001

17-JAN-2001; 2001WO-US01354

31-JAN-2000; 2000US-0179065.

24-FEB-2000; 2000US-0184664.

16-MAR-2000; 2000US-0189874.

17-MAR-2000; 2000US-0190076.



CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting  
CC the nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AA64703  
CC to AA87694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AA54942 to AA54950 and AA82169  
CC represent sequences used in the exemplification of the present invention.  
XX

SO Sequence 136 BP; 37 A; 34 C; 41 G; 24 T; 0 other;

## Alignment Scores:

Pred. No.:	0.137	Length:	136
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AA82863 (1-136)

QY 10 SerLeuGlyaspSerGluThrLeuSerGln 19

DB 99 AGCCTGGGTGACAGTGCACCTGCTCA 128

## RESULT 13

AAK82864

ID AAK82864 standard; DNA; 136 BP.

AC AAK82864;

XX 07-NOV-2001 (first entry)

DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:37676.

KM Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

OS WO200157182-A2.

PD 09-AUG-2001.

PE 17-JAN-2001; 2001WO-US01354.

XX 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 07-JUN-2000; 2000US-0205515.

PR 28-JUN-2000; 2000US-0209467.

PR 30-JUN-2000; 2000US-0214886.

PR 07-JUL-2000; 2000US-0215135.

PR 11-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0217496.

PR 26-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225457.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226686.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 05-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 06-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 08-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 12-SEP-2000; 2000US-0232081.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 14-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 21-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234997.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235484.

PR 27-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 02-OCT-2000; 2000US-0236370.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 13-OCT-2000; 2000US-0239937.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241786.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241809.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR	08-NOV-2000;	2000US-0246524.
PR	08-NOV-2000;	2000US-0246525.
PR	08-NOV-2000;	2000US-0246526.
PR	08-NOV-2000;	2000US-0246527.
PR	08-NOV-2000;	2000US-0246528.
PR	08-NOV-2000;	2000US-0246532.
PR	08-NOV-2000;	2000US-0246509.
PR	08-NOV-2000;	2000US-0246510.
PR	08-NOV-2000;	2000US-0246511.
PR	08-NOV-2000;	2000US-0246513.
PR	17-NOV-2000;	2000US-0249207.
PR	17-NOV-2000;	2000US-0249208.
PR	17-NOV-2000;	2000US-0249209.
PR	17-NOV-2000;	2000US-0249210.
PR	17-NOV-2000;	2000US-0249211.
PR	17-NOV-2000;	2000US-0249212.
PR	17-NOV-2000;	2000US-0249213.
PR	17-NOV-2000;	2000US-0249214.
PR	17-NOV-2000;	2000US-0249215.
PR	17-NOV-2000;	2000US-0249216.
PR	17-NOV-2000;	2000US-0249217.
PR	17-NOV-2000;	2000US-0249218.
PR	17-NOV-2000;	2000US-0249244.
PR	17-NOV-2000;	2000US-0249245.
PR	17-NOV-2000;	2000US-0249264.
PR	17-NOV-2000;	2000US-0249265.
PR	17-NOV-2000;	2000US-0249297.
PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250391.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	08-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
PI	Rosen CA, Barash SC, Ruben SM;	
XX		
DR	WPI: 2001-483426/52.	
XX		
PI	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PS	metastasis -	
XX		
XX	Disclosure: SEQ ID NO 37676; 3071pp + Sequence Listing; English.	
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting	
CC	the nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/hematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
XX	represent sequences used in the exemplification of the present invention.	

```

SQ Sequence 136 BP; 37 A; 34 C; 41 G; 24 T; 0 other;

Alignment Scores:
Pred. No.: 0.137 Length: 136
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
DB: 22 Gaps: 0

US-09-854-133-586 (1-97) x AAK82864 (1-136)
OY 10 SerleugIYASPSerGIuThrluSergIn 19
   |||||
Db 99 AGCGTGGGTGACAGTGAAGACACTCTCTCA 128

RESULT 14
AAC23449/C
ID AAC23449 standard; cDNA; 207 BP.
XX AAC23449;
AC
XX
XX AAC23449;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 27524.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
OS
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-0200610.
XX
XX 26-FEB-1999; 99US-0122487.
XX
XX (GEST ) GENSET.
XX
XX Dunas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
XX Claim 1; SEQ ID 27524; 71pp + CD-ROM; English.
XX
XX The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. No ORF has yet been conclusively
XX identified within the present sequence. The 5' ESTs were prepared from
XX total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
XX sequences usually correspond mainly to the 3' untranslated region (UTR)
XX of the mRNA because they are often obtained from oligo-dT primed cDNA
XX libraries. Such ESTs are not well suited for isolating cDNA sequences
XX derived from the 5' ends of mRNAs and even in those cases where longer
XX cDNA sequences have been obtained, the full 5' UTR is rarely included.
XX 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
XX used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
XX in diagnostic, forensic, gene therapy and chromosome mapping procedures
XX They are used to obtain upstream regulatory sequences and to design
XX expression and secretion vectors.
XX
XX Sequence 207 BP; 46 A; 44 C; 53 G; 60 T; 4 other;

Alignment Scores:
Pred. No.: 0.202 Length: 207
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

```

Query Match: 10.31% Indels: 0  
DB: 21 Gaps: 0

US-09-854-133-586 (1-97) x AAC23449 (1-207)

QY 10 SerLeuGLYASpSerGIurhLeuSerGln 19  
|||||  
DB 102 AGCCTGGGTGACAGTGAACCCCTGTCTCA 73

RESULT 15  
AAAA4648/c  
ID AAA44648 standard; cDNA: 291 BP.  
XX  
AC AAA44648;  
XX  
DT 21-AUG-2000 (first entry)  
XX  
DE Human secreted expressed sequence tag SEQ ID NO:1223.  
XX  
KW Human: mouse; chicken; rat; secreted expressed sequence tag; SESTR;  
KW expressed sequence tag; EST; probe; chemotactic; proliferative;  
KW immunomodulatory; haematopoietic; chemokine; analgesic; hemostatic;  
KW thrombolytic; antiinflammatory; cytostatic; antibacterial; antitumor;  
KW antiviral; antidiabetic; antistatic; vulnery; antiparkinsonian;  
KW anticancer; osteoprotective; neuroprotective; nontoxic; antiproliferative;  
KW vaccine; autoimmune disorder; multiple sclerosis; allergic condition;  
KW insulin dependent diabetes; asthma; myeloid cell deficiency; ulcer;  
KW lymphoid cell deficiency; burn; osteoporosis; osteoarthritis;  
KW central nervous system disorder; Alzheimer's disease; stroke;  
KW Parkinson's disease; Huntington's disease; coagulation disorder;  
KW haemophilia; thrombosis; inflammatory disorder; Crohn's disease;  
KW tumour; infection; depression; psoriasis; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200021991-A1.  
XX  
PD 20-APR-2000.  
XX  
PF 15-OCT-1999; 99WO-US24206.  
XX  
PR 15-OCT-1998; 98US-0104436.  
XX  
PA (GENY) GENETICS INST INC.  
XX  
PI Jacobs K, McCoy JM, LaVallie ER, Collins-Racie LA, Evans C;  
PI Merberg D, Treacy M, Bowman MR;  
XX  
DR WPI: 2000-317938/27.  
XX  
PT Isolated polynucleotides, and encoded proteins, comprising secreted  
PT expressed sequence tags (SESTs), useful for treating various disorders  
PT such as autoimmune, infectious, and central nervous system disorders -  
XX  
XX Claim 1; Page 516; 803pp; English.

AAAA43426 to AAA45925 represent specifically claimed expressed  
sequence tags (SESTs), isolated from human, mouse, chicken and rat  
tissue sources. The SESTs can have a range of activities depending on  
the tissues they were isolated from. The activities include:  
chemotactic; proliferative; immunomodulatory; haematopoietic;  
CC cytostatic; analgesic; haemostatic; thrombolytic; antiinflammatory;  
CC antistatic; antibacterial; antitumor; antidiabetic;  
CC nontoxic; antiparkinsonian; antiproliferative; neuroprotective;  
CC anticancer; osteoprotective; vulnery; antiparkinsonian;  
CC therapy and in vaccines. The SESTs can be used for gene  
identification and isolation of full-length cDNAs and genomic DNA  
molecules which correspond to the SESTs. Proteins encoded by the SESTs  
are useful in assays for determining biological activity and raising  
antibodies. They may be useful for treatment of autoimmune disorders  
(multiple sclerosis, insulin dependent diabetes), allergic conditions

CC (asthma), myeloid or lymphoid cell deficiencies, wounds, burns, ulcers,  
CC osteoporosis, osteoarthritis, central nervous system disorders,  
CC (Alzheimer's, Parkinson's, Huntington's disease, stroke), coagulation  
CC disorders (haemophilia, thrombosis), inflammatory disorders (Crohn's  
CC disease), tumours, bacterial, fungal or viral infections, depression and  
CC psoriasis. AAA45926 to AAA45931 represent linker variants which are given  
XX in the exemplification of the present invention.

SQ Sequence 291 BP; 59 A; 83 C; 57 G; 92 T; 0 other;

Alignment Scores:  
Pred. No.: 0.277 Length: 291  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 21 Gaps: 0

US-09-854-133-586 (1-97) x AAA44648 (1-291)

QY 10 SerLeuGLYASpSerGIurhLeuSerGln 19  
|||||  
DB 171 AGCCTGGGTGACAGTGAACCCCTGTCTCA 142

RESULT 16  
AAS37013  
ID AAS37013 standard; cDNA: 291 BP.  
XX  
AC AAS37013;  
XX  
DT 17-DEC-2001 (first entry)  
XX  
DE Novel human diagnostic and therapeutic gene #71.  
XX  
KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200166753-A2.  
XX  
PD 13-SEP-2001.  
XX  
PF 09-MAR-2001; 2001WO-US07787.  
XX  
PR 09-MAR-2000; 2000US-0188609.  
XX  
PA (CHIR) CHIRON CORP.  
PA (HYSE-) HYSEQ INC.  
XX  
PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;  
PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kaasam A, Lamson G;  
PI Drmanac R, Crkvenjakov R, Dickson M, Drmanac S, Labat I;  
PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;  
XX  
DR WPI: 2001-530177/58.  
XX  
PT New polynucleotides and polypeptides, useful for diagnosis and  
PT treatment of breast, lung and colon cancer -  
XX  
XX Claim 1; Page 613; 1193pp; English.

The invention relates to new polynucleotides and polypeptides, useful for  
diagnosis and treatment of breast, lung and colon cancer. The sequences  
can be used in detecting differentially expressed genes correlated with a  
cancerous state of a mammalian cell, comprising detecting at least one  
cell differentially expressed gene product in a test sample derived from a  
tumour suspected of being cancerous. They can also be used to inhibit  
cell growth by modulating expression of a gene product. AAS36943-  
AAS39338 represent novel human diagnostic and therapeutic coding  
sequences of the invention.

SQ Sequence 291 BP; 76 A; 61 C; 80 G; 74 T; 0 other;

## Alignment Scores:

Pred. No.: 0.277 Length: 291  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 22 Gaps: 0

US-09-854-133-586 (1-97) x AAS37013 (1-291)

QY 10 SerleuGlyAspSerGluThrLeuSerGln 19

DB 97 AGCCTGGTGACGACGAGACTCTATCTCAA 126

RESULT 17

AAS37176

ID AAS37176 standard; cDNA; 291 BP.

AC AAS37176;

DT 17-DEC-2001 (first entry)

XX Novel human diagnostic and therapeutic gene #234.

KW Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.

XX Homo sapiens.

PN WO200166753-A2.

XX 13-SEP-2001.

PF 09-MAR-2001; 2001WO-US07787.

PR 09-MAR-2000; 2000US-0188609.

XX (CHIR) CHIRON CORP.

PA (HYSE-) HYSEQ INC.

PI Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

PI Reinhard C, Randazzo F, Kennedy GC, Pot D, Kassam A, Lamson G;

PI Drmanac R, Crivenjakov R, Dickson M, Drmanac S, Labat I;

PI Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;

XX WPI; 2001-530177/58.

XX New polynucleotides and polypeptides, useful for diagnosis and

PT treatment of breast, lung and colon cancer -

XX Claim 1; Page 655; 1193pp; English.

PS The invention relates to new polynucleotides and polypeptides, useful for

CC diagnosis and treatment of breast, lung and colon cancer. The sequences

CC can be used in detecting differentially expressed genes correlated with a

CC cancerous state of a mammalian cell, comprising detecting at least one

CC differentially expressed gene product in a test sample derived from a

CC cell suspected of being cancerous. They can also be used to inhibit

CC tumour growth by modulating expression of a gene product. AAS36943-

CC AAS39338 represent novel human diagnostic and therapeutic coding

XX sequences of the invention.

XX SQ Sequence 291 BP; 74 A; 69 C; 77 G; 71 T; 0 other;

XX Alignment Scores:

Pred. No.: 0.277 Length: 291

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0

DB: 22 Gaps: 0

US-09-854-133-586 (1-97) x AAS37176 (1-291)

QY 10 SerleuGlyAspSerGluThrLeuSerGln 19

DB 28 AGCCTGGTGACGACGAGACTCTATCTCAA 57

RESULT 18

AAK75253

ID AAK75253 standard; DNA; 299 BP.

AC AAK75253;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:30065.

KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;

XX cytostatic; gene therapy; vaccine; metastasis; ds.

OS Homo sapiens.

PN WO200157182-A2.

XX 09-AUG-2001.

PF 17-JAN-2001; 2001WO-US01354.

PR 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0189874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 07-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 11-JUL-2000; 2000US-0217496.

PR 14-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 26-JUL-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0224519.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225266.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226279.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226688.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 05-SEP-2000; 2000US-0229513.

PR 06-SEP-2000; 2000US-0230437.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR	17-NOV-2000;	2000US-0249299.
PR	17-NOV-2000;	2000US-0249300.
PR	01-DEC-2000;	2000US-0250160.
PR	01-DEC-2000;	2000US-0250160.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251030.
PR	05-DEC-2000;	2000US-0251988.
PR	05-DEC-2000;	2000US-0256719.
PR	06-DEC-2000;	2000US-0251479.
PR	08-DEC-2000;	2000US-0251856.
PR	08-DEC-2000;	2000US-0251868.
PR	08-DEC-2000;	2000US-0251869.
PR	08-DEC-2000;	2000US-0251989.
PR	11-DEC-2000;	2000US-0251990.
PR	11-DEC-2000;	2000US-0254097.
PR	05-JAN-2001;	2001US-0259678.
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-483426/52.	
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and	
PT	metastasis -	
PS	Disclosure; SEQ ID NO. 30065; 3071pp + Sequence Listing; English.	
XX		
CC	AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)	
CC	amino acid sequences given in AAM82170 to AAM91921. (I) have cytosolic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patients' own production or (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting	
CC	the nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/hematopoietic-related diseases, especially	
CC	cancers and cancer metastases of hematopoietic-derived cells. AAK64703	
CC	to AAK81694 represent human immune/hematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAM82169	
CC	represent sequences used in the exemplification of the present invention.	
XX		
SQ	Sequence 299 BP; 99 A; 71 C; 80 G; 49 T; 0 other;	
Alignment Scores:		
Pred. No.:	0.285	Length: 299
Score:	10.00	Matches: 10
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	10.31%	Indels: 0
BB:	22	Gaps: 0
US-09-854-133-586 (1-97) x AAK75253 (1-299)		
OY	10 SerLeuGIyAspSerGluThrLeuSergin 19	
Db	AAAAAAAAAAAAAAAAAAAAA	
235 AGCCTGGCGACAGTGAAGACTGTGTCATA 264		
RESULT 19		
AAAZ14169		
ID	AAZ14169 standard; cDNA: 300 BP.	
XX		
AC	AAZ14169;	
XX		
DT	12-OCT-1999 (first entry)	
XX		
DE	Human gene expression product cDNA sequence SEQ ID NO:1638.	
XX		
TW	Human; gene; gene expression product; diagnosis; therapy. probe.	



KW detection: mapping; tissue typing; profiling; forensic; cancer;  
 KM genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
 OS Homo sapiens.  
 XX  
 PN WO9938972-A2.  
 XX  
 PD 05-AUG-1999.  
 XX  
 PF 28-JAN-1999; 99WO-US01619.  
 XX  
 PR 03-APR-1998; 98US-0080666.  
 PR 28-JAN-1998; 98US-0072910.  
 PR 24-FEB-1998; 98US-0075954.  
 PR 31-MAR-1998; 98US-0080114.  
 PR 03-APR-1998; 98US-0080515.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX (HYSE-) HYSEQ INC.  
 XX  
 PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
 PI Escobedo J, Garcia PD, Garcia V, Glese R, Innis MA;  
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
 XX  
 DR WPI: 1999-494092/41.  
 XX  
 PT Novel human genes and their expression products which are  
 PT differentially expressed in different cell types  
 XX  
 PS Claim 1; Page 1031; 2479pp; English.  
 XX  
 CC The present invention describes a library of human polynucleotides  
 CC comprising the sequences given in AA21352 to AA21779. Also described is  
 CC a method of detecting differentially expressed genes correlated with the  
 CC cancerous state of a mammalian cell, comprising detecting at least one  
 CC differentially expressed gene product in a test sample from a cell  
 CC suspected of being cancerous, where the gene product is encoded by one  
 CC of the 5248 polynucleotide sequences given in AA21352 to AA21779. The  
 CC polynucleotides can be used as a source of primers and probes, which can  
 CC be used for a variety of purpose, e.g. detection of expression levels,  
 CC mapping, tissue typing or profiling, forensics, genetic analysis and  
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
 CC can be used for raising antibodies for experimental, diagnostic and  
 CC therapeutic purposes. The polynucleotides may also be used to construct  
 CC arrays for diagnostics (which may be used to determine function of an  
 CC encoded protein); and to detect differences in expression levels between  
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
 CC identify a genetic predisposition or susceptibility to a disease such as  
 CC cancer). The polynucleotides of the invention are especially used in the  
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
 CC and lung cancer. The polynucleotides can also be used to screen for  
 CC peptide analogues and antagonists.  
 XX  
 SQ Sequence 300 BP; 86 A; 63 C; 77 G; 74 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0.285 Length: 300  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 20 Gaps: 0  
 US-09-854-133-586 (1-97) x AA214169 (1-300)  
 OY 61 GluserleuLeucysProPseProlys 70  
 DB 270 GAAAGCCTGTGTGTCACCATCTCCAAAG 299  
 RESULT 20  
 AA213802/c

ID AA213802 standard; cDNA; 300 BP.  
 XX AA213802;  
 AC  
 XX  
 DT 12-OCT-1999 (first entry)  
 XX  
 DE Human gene expression product cDNA sequence SEQ ID NO:1271.  
 XX  
 KW Human; gene; gene expression product; diagnosis; therapy; probe;  
 KW detection; mapping; tissue typing; profiling; forensic; cancer;  
 KW genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9938972-A2.  
 XX  
 PD 05-AUG-1999.  
 XX  
 PF 28-JAN-1999; 99WO-US01619.  
 XX  
 PR 03-APR-1998; 98US-0080666.  
 PR 28-JAN-1998; 98US-0072910.  
 PR 24-FEB-1998; 98US-0075954.  
 PR 31-MAR-1998; 98US-0080114.  
 PR 03-APR-1998; 98US-0080515.  
 XX  
 PA (CHIR ) CHIRON CORP.  
 XX (HYSE-) HYSEQ INC.  
 XX  
 PI Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;  
 PI Escobedo J, Garcia PD, Garcia V, Glese R, Innis MA;  
 PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;  
 PI Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;  
 PI Stache-Crain B, Sudduth-Klinger J, Williams LT;  
 XX  
 DR WPI: 1999-494092/41.  
 XX  
 PT Novel human genes and their expression products which are  
 PT differentially expressed in different cell types  
 XX  
 PS Claim 1; Page 950; 2479pp; English.  
 XX  
 CC The present invention describes a library of human polynucleotides  
 CC comprising the sequences given in AA21352 to AA21779. Also described is  
 CC a method of detecting differentially expressed genes correlated with the  
 CC cancerous state of a mammalian cell, comprising detecting at least one  
 CC differentially expressed gene product in a test sample from a cell  
 CC suspected of being cancerous, where the gene product is encoded by one  
 CC of the 5248 polynucleotide sequences given in AA21352 to AA21779. The  
 CC polynucleotides can be used as a source of primers and probes, which can  
 CC be used for a variety of purpose, e.g. detection of expression levels,  
 CC mapping, tissue typing or profiling, forensics, genetic analysis and  
 CC detection of polymorphisms. Polypeptides encoded by the polynucleotides  
 CC can be used for raising antibodies for experimental, diagnostic and  
 CC therapeutic purposes. The polynucleotides may also be used to construct  
 CC arrays for diagnostics (which may be used to determine function of an  
 CC encoded protein); and to detect differences in expression levels between  
 CC two cells (e.g. to identify abnormal or diseased tissue in a human, to  
 CC identify a genetic predisposition or susceptibility to a disease such as  
 CC cancer). The polynucleotides of the invention are especially used in the  
 CC diagnosis, prognosis and management of colorectal cancer, breast cancer,  
 CC and lung cancer. The polynucleotides can also be used to screen for  
 CC peptide analogues and antagonists.  
 XX  
 SQ Sequence 300 BP; 62 A; 83 C; 62 G; 93 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0.285 Length: 300  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 20 Gaps: 0

US-09-854-133-586 (1-97) x AA213802 (1-300)

OY 10 SerLeuGlyAspSerGluThrLeuSerGln 19

Db 131 AGCCTGGGTGACAGTGTCTCTCA 102

RESULT 21

AA213190

ID AA213190 standard; CDNA: 300 BP.

AA213190;

12-OCT-1999 (first entry)

Human gene expression product cDNA sequence SEQ ID NO:659.

Human: gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.

Homo sapiens.

MO9938972-A2.

05-AUG-1999.

28-JAN-1999; 99WO-US01619.

03-APR-1998; 98US-0080666.

28-JAN-1998; 98US-0072810.

24-FEB-1998; 98US-0075954.

31-MAR-1998; 98US-0080114.

03-APR-1998; 98US-0080515.

(CHIR) CHIRON CORP.

(HYSE-) HYSEQ INC.

Escobedo J, Garcia PD, Garcia V, Glese K, Innis MA;

Jones WL, Kaasam A, Kennedy GC, Kita D, Labat I;

Lamson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;

Stache-Crain B, Sudduth-Klinger J, Williams LT;

WPI: 1999-494092/41.

Novel human genes and their expression products which are

differentially expressed in different cell types

Claim 1; Page 813; 2479pp; English.

The present invention describes a library of human polynucleotides

comprising the sequences given in AA212532 to AA21779. Also described is

a method of detecting differentially expressed genes correlated with the

cancerous state of a mammalian cell, comprising detecting at least one

differentially expressed gene product in a test sample from a cell

suspected of being cancerous, where the gene product is encoded by one

of the 548 polynucleotide sequences given in AA212532 to AA21779. The

polynucleotides can be used as a source of primers and probes, which can

be used for a variety of purpose, e.g. detection of expression levels,

mapping, tissue typing or profiling, forensics, genetic analysis and

detection of polymorphisms. Polypeptides encoded by the polynucleotides

can be used for raising antibodies for experimental, diagnostic and

therapeutic purposes. The polynucleotides may also be used to construct

arrays for diagnosis; and to detect differences in expression levels between

two cells (e.g. to identify abnormal or diseased tissue in a human, to

identify a genetic predisposition or susceptibility to a disease such as

cancer). The polynucleotides of the invention are especially used in the

diagnosis, prognosis and management of colorectal cancer, breast cancer,

and lung cancer. The polynucleotides can also be used to screen for

peptide analogues and antagonists.

Sequence 300 BP; 74 A; 72 C; 86 G; 68 T; 0 other;

Alignment Scores:

Pred. No.: 0.285 Length: 300

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0

DB: 20 Gaps: 0

US-09-854-133-586 (1-97) x AA213190 (1-300)

OY 10 SerLeuGlyAspSerGluThrLeuSerGln 19

Db 262 AGCCTGGGTGACAGTGTCTCTCA 291

RESULT 22

AA2137025

ID AA2137025 standard; CDNA: 303 BP.

AA2137025;

17-DEC-2001 (first entry)

Novel human diagnostic and therapeutic gene #83.

Human; cancer; breast; lung; colon; prostate; cytostatic; diagnostic; ss.

Homo sapiens.

WO200166753-A2.

13-SEP-2001.

09-MAR-2001; 2001WO-US07787.

09-MAR-2000; 2000US-0188609.

(CHIR) CHIRON CORP.

(HYSE-) HYSEQ INC.

Williams LT, Escobedo J, Innis MA, Garcia PD, Sudduth-Klinger J;

Reinhard C, Randazzo F, Kennedy GC, Pot D, Kaasam A, Lamson G;

Drmannac R, Crivenjakov R, Dickson M, Drmanac S, Labat I;

Leshkowitz D, Kita D, Garcia V, Jones WL, Stache-Crain B;

WPI: 2001-530177/58.

Novel polynucleotides and polypeptides, useful for diagnosis and

treatment of breast, lung and colon cancer -

Claim 1; Page 616; 1193pp; English.

The invention relates to new polynucleotides and polypeptides, useful for

diagnosis and treatment of breast, lung and colon cancer. The sequences

can be used in detecting differentially expressed genes correlated with a

cancerous state of a mammalian cell, comprising detecting at least one

differentially expressed gene product in a test sample derived from a

cell suspected of being cancerous. They can also be used to inhibit

tumour growth by modulating expression of a gene product. AA2136943-

sequences of the invention.

Sequence 303 BP; 79 A; 65 C; 84 G; 75 T; 0 other;

Alignment Scores:

Pred. No.: 0.288 Length: 303

Score: 10.00 Matches: 10

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0

DB: 22 Gaps: 0

US-09-854-133-586 (1-97) x AAS37025 (1-303)  
QY 10 SerLeuGIyAspSerGIuThrLeuSerGIu 19  
|||||  
Db 97 AGCCTGGGTGACACGAGACTCTATCTCA 126  
RESULT 23  
AAK75254  
ID AAK75254 standard; DNA; 310 BP.  
XX  
AC AAK75254;  
XX  
DT 07-NOV-2001 (first entry)  
XX  
DE Human Immune/haematopoietic antigen genomic sequence SEQ ID NO:30066.  
XX  
KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US01354.  
XX  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 11-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225278.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226686.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229309.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234597.  
PR 25-SEP-2000; 2000US-0234598.  
PR 26-SEP-2000; 2000US-0234584.  
PR 27-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251889.  
 PR 11-DEC-2000; 2000US-0251890.  
 PR 05-JAN-2001; 2001US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 DR WPI; 2001-483426/52.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 PT useful for preventing, diagnosing and/or treating cancers and  
 PT metastasis -  
 PS Disclosure: SEQ ID NO 30066; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK7664 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
 CC represent sequences used in the exemplification of the present invention.  
 CC  
 SQ Sequence 310 BP; 97 A; 75 C; 90 G; 48 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.294 Length: 310  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-854-133-586 (1-97) x AAK75254 (1-310)  
 QY 10 SerLeuGIyASpSerGIuThrlauSerGIu 19  
 Db 252 AGCCTGGCGACAGTACGACTGTCTCTCA 281  
 RESULT 24  
 AAI92003  
 ID AAI92003 standard; cDNA; 343 BP.  
 XX  
 AC AAI92003;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 12063.

XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PE 26-FEB-2001; 2001WO-US04927.  
 XX  
 PR 28-FEB-2000; 2000US-0515126.  
 PR 18-MAY-2000; 2000US-0577409.  
 XX  
 PA (HYSE-) HYSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI; 2001-514838/56.  
 DR P-PDB; AAO12072.  
 XX  
 PT Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosing and treating e.g. leukaemia, inflammation and immune  
 PT disorders -  
 PS Claim 1; SEQ ID NO 12063; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pat\_sequences.  
 CC  
 SQ Sequence 343 BP; 85 A; 87 C; 90 G; 81 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.: 0.323 Length: 343  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-854-133-586 (1-97) x AAI92003 (1-343)  
 QY 10 SerLeuGIyASpSerGIuThrlauSerGIu 19  
 Db 94 AGCCTGGCGACAGCAGACTGTCTCTCA 123  
 RESULT 25  
 ABL6133/C  
 ID ABL6133 standard; DNA; 351 BP.  
 XX  
 AC ABL6133;  
 XX  
 DT 15-MAY-2002 (first entry)  
 XX  
 DE Breast cancer related gene sequence SEQ ID NO:1470.  
 XX  
 KW Human; cancer; colon; breast; oesophagus; kidney; thyroid;  
 KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;

KM cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
 KM gene; ds.  
 XX Homo sapiens.  
 OS  
 PN WO200194629-A2.  
 XX  
 PD 13-DEC-2001.  
 XX  
 PF 30-MAY-2001; 2001WO-US10838.  
 XX  
 PR 05-JUN-2000; 2000US-209473P.  
 PR 18-SEP-2000; 2000US-209531P.  
 PR 18-SEP-2000; 2000US-233133P.  
 PR 18-SEP-2000; 2000US-233617P.  
 PR 20-SEP-2000; 2000US-234009P.  
 PR 20-SEP-2000; 2000US-234034P.  
 PR 20-SEP-2000; 2000US-234052P.  
 PR 22-SEP-2000; 2000US-234509P.  
 PR 22-SEP-2000; 2000US-234567P.  
 PR 25-SEP-2000; 2000US-234923P.  
 PR 25-SEP-2000; 2000US-234924P.  
 PR 25-SEP-2000; 2000US-235077P.  
 PR 25-SEP-2000; 2000US-235082P.  
 PR 25-SEP-2000; 2000US-235134P.  
 PR 26-SEP-2000; 2000US-235637P.  
 PR 26-SEP-2000; 2000US-235638P.  
 PR 27-SEP-2000; 2000US-235711P.  
 PR 27-SEP-2000; 2000US-235720P.  
 PR 27-SEP-2000; 2000US-235840P.  
 PR 28-SEP-2000; 2000US-235863P.  
 PR 28-SEP-2000; 2000US-236028P.  
 PR 28-SEP-2000; 2000US-236032P.  
 PR 28-SEP-2000; 2000US-236033P.  
 PR 28-SEP-2000; 2000US-236034P.  
 PR 28-SEP-2000; 2000US-236109P.  
 PR 28-SEP-2000; 2000US-236111P.  
 PR 29-SEP-2000; 2000US-236842P.  
 PR 29-SEP-2000; 2000US-236891P.  
 PR 02-OCT-2000; 2000US-237172P.  
 PR 02-OCT-2000; 2000US-237173P.  
 PR 02-OCT-2000; 2000US-237278P.  
 PR 02-OCT-2000; 2000US-237294P.  
 PR 02-OCT-2000; 2000US-237316P.  
 PR 02-OCT-2000; 2000US-237319P.  
 PR 03-OCT-2000; 2000US-237425P.  
 PR 03-OCT-2000; 2000US-237598P.  
 PR 03-OCT-2000; 2000US-237604P.  
 PR 03-OCT-2000; 2000US-237606P.  
 PR 03-OCT-2000; 2000US-237608P.  
 PR 01-NOV-2000; 2000US-244867P.  
 PR 01-NOV-2000; 2000US-245084P.  
 XX  
 PA (AVAL-) AVALON PHARM.  
 XX  
 PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
 PI Soppe DR, Weaver Z;  
 XX WPI; 2002-188264/24.  
 DR  
 XX  
 XX  
 PT Screening for anti-neoplastic agent involves exposing cells to a  
 PT chemical agent to be tested for anti-neoplastic activity, and  
 PT determining a change in expression of a gene of a signature gene set -  
 XX  
 PS Claim 1; SEQ ID 1470; 44pp; English.  
 XX  
 CC The present invention describes a method (M1) for screening for an  
 CC anti-neoplastic agent. The method involves exposing cells to a chemical  
 CC agent to be tested for anti-neoplastic activity, determining a change in  
 CC expression of at least one gene (I) of a signature gene set, where (I)  
 CC comprises a sequence (s) selected from 8447 sequences (given in ABU61664  
 CC to ABU70110), or is at least 95% identical to (S), where a change in

CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
 CC activity and can be used in gene therapy. M1 can be used for screening  
 CC an anti-neoplastic agent, and can be used for producing a product which  
 CC is the data collected with respect to the anti-neoplastic agent as a  
 CC result of M1, and the data is sufficient to convey the chemical  
 CC structure and/or properties of the agent. M1 can be used in the  
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
 CC carcinoma, papillary carcinoma and Wilm's tumour.  
 XX  
 SQ Sequence 351 BP; 88 A; 92 C; 67 G; 101 T; 3 other;  
 Alignment Scores:  
 Pred. No.: 0.33 Length: 351  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 24 Gaps: 0  
 US-09-854-133-586 (1-97) x ABU63133 (1-351)  
 OY 10 SerLeuGlyaspSergIuThrIeuSergIn 19  
 DB 153 AGCCTGGGTGACAGCGAGACCCCTCTCA 124  
 RESULT 26  
 ABV19284  
 ID ABV19284 standard; cDNA; 371 BP.  
 XX  
 AC ABV19284;  
 XX  
 DT 13-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker cDNA 19275.  
 XX  
 DE Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KM pharmacogenomic marker; gene; ss.  
 OS  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 PR 16-MAR-2000; 2000US-189862P.  
 PR 25-MAY-2000; 2000US-207454P.  
 PR 09-JUN-2000; 2000US-211314P.  
 PR 18-JUL-2000; 2000US-219007P.  
 PR 13-DEC-2000; 2000US-255281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI; 2001-662795/76.  
 XX  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer. Useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX  
 PS Claim 1; Page 3158; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 371 BP; 112 A; 87 C; 87 G; 85 T; 0 other;

Alignment Scores:  
Pred. No.: 0.348 Length: 371  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV19284 (1-371)

QY 10 Serleuglyaspserglurthrlausergln 19  
|||||  
DB 223 AGCCTGGGCGACAGTGAACCTGTCTCA 252

RESULT 27  
ABV13690  
ID ABV13690 standard; cDNA; 376 BP.  
XX  
AC ABV13690;  
XX  
DT 13-SEP-2002 (first entry)  
XX  
DE Human prostate expression marker cDNA 13681.  
XX  
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200160860-A2.  
XX  
PD 23-AUG-2001.  
XX  
PF 20-FEB-2001; 2001WO-US05171.  
XX  
PR 17-FEB-2000; 2000US-183319P.  
PR 16-MAR-2000; 2000US-189862P.  
PR 25-MAY-2000; 2000US-207454P.  
PR 09-JUN-2000; 2000US-211314P.  
PR 18-JUL-2000; 2000US-219007P.  
PR 13-DEC-2000; 2000US-255281P.  
XX  
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
XX  
PI Schlegel R, Endege WO, Monahan JE;  
XX  
DR WPI; 2001-662795/76.  
XX  
PT Novel isolated nucleic acid molecule associated with cancerous state of  
PT prostate cells and correlating with presence of prostate cancer, useful  
PT for detecting presence of prostate cancer, stage of prostate cancer -  
XX  
PS Claim 1; Page 2271-2272; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (i) comprising  
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
CC specification or its complement. (i) is useful for:  
CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 376 BP; 102 A; 68 C; 103 G; 103 T; 0 other;

Alignment Scores:  
Pred. No.: 0.352 Length: 376  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV13690 (1-376)

QY 10 Serleuglyaspserglurthrlausergln 19  
|||||  
DB 57 AGCCTGGGCGACAGTGAACCTGTCTCA 86

RESULT 28  
AAI89561  
ID AAI89561 standard; cDNA; 394 BP.  
XX  
AC AAI89561;  
XX  
DT 06-NOV-2001 (first entry)  
XX  
DE Human polynucleotide SEQ ID NO 9621.  
XX  
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; hematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukemia;  
KW nervous system disorders; arthritis; inflammation; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200164835-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 26-FEB-2001; 2001WO-US04927.  
XX  
PR 28-FEB-2000; 2000US-0515126.  
PR 18-MAY-2000; 2000US-0577409.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Tang YT, Liu C, Drmanac RT;  
XX  
DR WPI; 2001-514838/56.  
XX  
PT P-PSDB; AAO09630.  
XX  
PT Isolated nucleic acids and polypeptides, useful for preventing  
PT diagnosing and treating e.g. leukemia, inflammation and immune  
PT disorders -  
XX  
PS Claim 1; SEQ ID NO 9621; 1399pp + Sequence listing; English.

The invention relates to human polynucleotides (AAI79941-AAI93841) and  
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, hematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 394 BP; 128 A; 69 C; 115 G; 82 T; 0 other;

Alignment Scores:

Pred. No.:	0.368	Length:	394
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AA189561 (1-394)

OY 10 SerLeuGlyAspSerGluThrLeuSerGln 19

DB 117 AGCCTGGCGCAGAGTGAGACCCCTGTCTCAA 146

RESULT 29

ABV04521  
ID ABV04521 standard; CDNA; 401 BP.

XX ABV04521;

DT 13-SEP-2002 (first entry)

DE Human prostate expression marker CDNA 4512.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX WO200160860-A2.

PN 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer -

PS Claim 1; Page 783; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;  
CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence 401 BP; 102 A; 77 C; 115 G; 104 T; 3 other;

Alignment Scores:

Pred. No.:	0.374	Length:	401
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV04521 (1-401)

OY 10 SerLeuGlyAspSerGluThrLeuSerGln 19

DB 81 AGCCTGGCGCAGACTGTGCTCAA 110

RESULT 30

ABV34803  
ID ABV34803 standard; CDNA; 401 BP.

XX ABV34803;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker CDNA 34794.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX WO200160860-A2.

PN 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of  
XX prostate cells and correlating with presence of prostate cancer, useful  
XX for detecting presence of prostate cancer, stage of prostate cancer -

PS Claim 1; Page 7283; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising  
XX a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
XX specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;  
CC (b) monitoring the progression of prostate cancer in a patient;  
CC (c) assessing the efficacy of a test compound to inhibit prostate  
CC cancer in a patient;  
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
CC in a patient;  
CC (e) selecting the prostate cell carcinogenic potential of a compound;  
CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;  
CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 401 BP; 104 A; 78 C; 114 G; 105 T; 0 other;

## Alignment Scores:

Pred. No.:	0.374	Length:	401
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV43654 (1-401)

QY 10 SerLeuGLYASpSerGIurThreUSerGIu 19  
Db 82 AGCCTGGGCGACAGTGAAGACTTTGTCTCA 111

## RESULT 31

ID ABV43654 standard; cDNA; 401 BP.  
XX  
AC ABV43654;  
XX

16-SEP-2002 (first entry)

Human prostate expression marker cDNA 43645.

Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

pharmacogenomic marker; gene; ss.

Homo sapiens.

WO200160860-A2.

23-AUG-2001.

20-FEB-2001; 2001WO-US05171.

17-FEB-2000; 2000US-183319P.

16-MAR-2000; 2000US-189862P.

25-MAY-2000; 2000US-207454P.

09-JUN-2000; 2000US-211314P.

18-JUL-2000; 2000US-219007P.

13-DEC-2000; 2000US-255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

Schlegel R, Endege WO, Monahan JE;

WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of  
prostate cells and correlating with presence of prostate cancer, useful  
for detecting presence of prostate cancer, stage of prostate cancer -

Claim 1: Page 8690; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising  
a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
specification or its complement. (I) is useful for:

(a) assessing whether a patient is afflicted with prostate cancer;

(b) monitoring the progression of prostate cancer in a patient;

(c) assessing the efficacy of a test compound to inhibit prostate  
cancer in a patient;(d) assessing the efficacy of a therapy for inhibiting prostate cancer  
in a patient;

(e) selecting a composition for inhibiting prostate cancer in a patient;

(f) assessing the prostate cell carcinogenic potential of a compound;

(g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
CC patient;  
CC (i) is also useful as a pharmacodynamic or pharmacogenomic marker.  
XX  
SQ Sequence 401 BP; 104 A; 78 C; 114 G; 105 T; 0 other;

## Alignment Scores:

Pred. No.:	0.374	Length:	401
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	23	Gaps:	0

US-09-854-133-586 (1-97) x ABV43654 (1-401)

QY 10 SerLeuGLYASpSerGIurThreUSerGIu 19  
Db 82 AGCCTGGGCGACAGTGAAGACTTTGTCTCA 111

## RESULT 32

ID ABL84142 standard; cDNA; 418 BP.  
XX  
AC ABL84142;  
XX

17-MAY-2002 (first entry)

Human ovarian cancer related cDNA clone SEQ ID NO:7120.

Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

Homo sapiens.

WO200192581-A2.

06-DEC-2001.

29-MAY-2001; 2001WO-US17756.

26-MAY-2000; 2000US-207484P.

(CORI-) CORIXA CORP.

Algate PA, Harlocker SL, Jones R;

WPI; 2002-122075/16.

Composition for therapy and diagnosis of ovarian cancer comprising  
polypeptide of a ovarian tumor polypeptide, polynucleotide encoding  
polypeptide, antibody specific to polypeptide or T cell expressing  
polypeptide -

Claim 1; SEQ ID 7120; 489pp; English.

The present invention describes a composition (I) comprising: carriers  
and immunostimulants; and a polypeptide (II) of a ovarian tumor  
polypeptide encoded by a polynucleotide (III) having a cDNA sequence  
(S1) from the 10912 nucleotide sequences as given in ABL77023 to  
ABL87934, (II) encoding (II) having a sequence (S2), a T cell  
population of (II), or antigen presenting cells that express (II).

(I) has cytostatic activity. An oligonucleotide (IV) that hybridises to  
(S1) can be used for detecting ovarian cancer in a patient's biological  
sample preferably serum or ovarian tissue. The method comprises  
contacting a biological sample from a patient with (IV), detecting the  
amount of polynucleotide hybridising to (IV) and comparing the amount to  
a predetermined cutoff value and thereby detecting ovarian cancer in the  
patient, where the amount of polynucleotide hybridising to (IV) isdetected preferably by polymerase chain reaction (PCR). (I) comprising  
(III) and/or (II) is useful for stimulating and/or expanding T cells  
specific for an ovarian tumour protein comprising contacting T cells  
with (III) or (II). (III) is useful in design and preparation of

ribozyme molecules for inhibiting expression of the tumour polypeptides



CC and proteins in tumour cells; and to isolate a full length gene from a  
 CC suitable library e.g., a tumour CDNA library using well known  
 CC techniques.  
 XX  
 SQ Sequence 418 BP; 93 A; 92 C; 77 G; 156 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.389 Length: 418  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 Gaps: 0  
 DB: 24  
 US-09-854-133-586 (1-97) x ABL84142 (1-418)  
 QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
 ||||||||||||||||||  
 Db 336 AGCCTGGGAGACAGACAGACCTGTCTCA 307  
 RESULT 33  
 ID AAI92526 standard; cDNA: 430 BP.  
 XX AAI92526;  
 AC AAI92526;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human polynucleotide SEQ ID NO 12586.  
 XX  
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorders; arthritis; inflammation; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200164835-A2.  
 XX  
 PD 07-SEP-2001.  
 XX  
 PF 26-FEB-2001; 2001WO-US04927.  
 XX  
 PR 28-FEB-2000; 2000US-0515126.  
 XX  
 PR 18-MAY-2000; 2000US-0577409.  
 XX  
 PA (HXSE-) HXSEQ INC.  
 XX  
 PI Tang YT, Liu C, Drmanac RT;  
 XX  
 DR WPI: 2001-514838/56.  
 XX  
 DR P-PSDB: AAO12595.  
 XX  
 PT Isolated nucleic acids and polypeptides, useful for preventing  
 PT diagnosis and treating e.g. leukaemia, inflammation and immune  
 PT disorders -  
 XX  
 PS Claim 1: SEQ ID NO 12586; 1399pp + Sequence Listing; English.  
 XX  
 CC The invention relates to human polynucleotides (AAI79941-AAI93841) and  
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activity/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pcl\_sequences.

XX  
 SQ Sequence 430 BP; 142 A; 78 C; 109 G; 101 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.399 Length: 430  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 Gaps: 0  
 DB: 22  
 US-09-854-133-586 (1-97) x AAI92526 (1-430)  
 QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
 ||||||||||||||||||  
 Db 104 AGCCTGGGTGACAGTGAACCCCTGTCTCA 133  
 RESULT 34  
 ID ABV49058 standard; cDNA: 440 BP.  
 XX ABV49058;  
 AC ABV49058;  
 XX  
 DT 17-SEP-2002 (first entry)  
 XX  
 DE Human prostate expression marker CDNA 49049.  
 XX  
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;  
 KW pharmacogenomic marker; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200160860-A2.  
 XX  
 PD 23-AUG-2001.  
 XX  
 PF 20-FEB-2001; 2001WO-US05171.  
 XX  
 PR 17-FEB-2000; 2000US-183319P.  
 XX  
 PR 16-MAR-2000; 2000US-189862P.  
 XX  
 PR 25-MAY-2000; 2000US-207454P.  
 XX  
 PR 09-JUN-2000; 2000US-211314P.  
 XX  
 PR 18-JUL-2000; 2000US-219007P.  
 XX  
 PR 13-DEC-2000; 2000US-235281P.  
 XX  
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Schlegel R, Endege WO, Monahan JE;  
 XX  
 DR WPI: 2001-662795/76.  
 XX  
 DR  
 XX  
 PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -  
 XX  
 PS Claim 1: Page 9597; 11750pp; English.  
 XX  
 CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.  
 XX

SQ Sequence 440 BP; 132 A; 107 C; 105 G; 96 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 0.408 Length: 440  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 23 Gaps: 0  
 US-09-854-133-586 (1-97) x ABV49058 (1-440)  
 QY 10 Serleuglyaspserglutthreusergl 19  
 Db 261 AGCCTGGCGACAGACGACCTGTCTCA 290  
 RESULT 35  
 AAH03381  
 ID AAH03381 standard; cDNA; 457 BP.  
 AC AAH03381;  
 XX  
 DT 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA clone (5'-primer) SEQ ID NO:216.  
 XX  
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 PN EP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 PR 29-JUL-1999; 99JP-0248036.  
 PR 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.  
 PR 02-MAY-2000; 2000JP-0183767.  
 PR 09-JUN-2000; 2000JP-0241899.  
 XX  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI; 2001-318749/34.  
 XX  
 PT Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 PS Claim 1; SEQ ID 216; 2537bp + CD ROM; English.  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC of an oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC complementary strand of a polynucleotide which complementary to the  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC in gene therapy. The primer sets can be used in antisense therapy and  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length

CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX  
 SQ Sequence 457 BP; 137 A; 91 C; 125 G; 101 T; 3 other;  
 Alignment Scores:  
 Pred. No.: 0.423 Length: 457  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-854-133-586 (1-97) x AAH03381 (1-457)  
 QY 10 Serleuglyaspserglutthreusergl 19  
 Db 393 AGCCTGGCGACAGACGACCTGTCTCA 422  
 RESULT 36  
 AA198951/C  
 ID AA198951 standard; DNA; 459 BP.  
 AC AA198951;  
 XX  
 DT 07-JAN-2002 (first entry)  
 XX  
 DE Human excretory related polynucleotide SEQ ID NO 715.  
 XX  
 KW Human; nocotropic; neuroprotective; cytosolic; dermatological; virucide;  
 KW immunosuppressive; anti-inflammatory; anti-HIV; antibacterial; vulnery;  
 KW antiparkinsonian; antischistosomal; antianemic; antidiabetic; cancer;  
 KW antipneumatic; hepatoprotective; cerebroprotective; anti-inflammatory;  
 KW antiparasitic; cardiant; antidiabetic; anticonvulsant; antifungal;  
 KW neurological disease; infection; immune disorder; cardiovascular disorder;  
 KW excretory system; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200155313-A2.  
 XX  
 PD 02-AUG-2001.  
 XX  
 PF 17-JAN-2001; 2001WO-US01323.  
 XX  
 PR 31-JAN-2000; 2000US-0179065.  
 PR 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 07-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 11-JUL-2000; 2000US-0217496.  
 PR 14-JUL-2000; 2000US-0218290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 26-JUL-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225216.  
 PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226629.  
 PR 22-AUG-2000; 2000US-0226661.  
 PR 22-AUG-2000; 2000US-0226688.  
 PR 22-AUG-2000; 2000US-0227182.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 01-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 05-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 06-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.  
 PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0231968.  
 PR 14-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234223.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 25-SEP-2000; 2000US-0234997.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 26-SEP-2000; 2000US-0235484.  
 PR 27-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239393.  
 PR 13-OCT-2000; 2000US-0239393.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0244674.  
 PR 08-NOV-2000; 2000US-0244675.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251859.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 PI WPI; 2001-465569/50.  
 XX  
 DR  
 XX  
 XX  
 PT Isolated nucleic acid molecule encoding excretory system antigen is  
 PT used in preventing, treating or ameliorating a medical condition -  
 XX  
 XX  
 PS Example 2; SEQ ID NO 715; 574bp + Sequence Listing; English.  
 XX  
 CC The invention relates to novel excretory system related human  
 CC polypeptides (AM98567-AM99503) and the encoded proteins  
 CC (AM99594-AM99913) useful for preventing, treating or ameliorating  
 CC medical conditions e.g. by protein or gene therapy, especially  
 CC disorders related to the excretory system. The genes are isolated  
 CC from a range of human tissues disclosed in the specification. The  
 CC nucleic acids, proteins, antibodies and (ant)agonists are useful in  
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast  
 CC and ovarian cancer and other cancers of the adrenal gland, bone, bone  
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;  
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune  
 CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's  
 CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative  
 CC colitis; (c) cardiovascular disorders such as myocardial ischaemia;  
 CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and  
 CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal  
 CC and parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 XX Sequence 459 BP; 106 A; 108 C; 123 G; 122 T; 0 other;  
 SQ

## Alignment Scores:

Pred. No.:	0.424	Length:	459
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAI98951 (1-459)

QY 10 SerLeuGLYAspSerGluhrhLeuSerGln 19

Db 373 AGCCTGGGTGACAGCGAGACCTACTCTCA 344

## RESULT 37

AAI6301/c

ID AAI63301 standard; DNA; 459 BP.

AC AAI63301;

XX 22-OCT-2001 (first entry)

DE Human kidney related polynucleotide SEQ ID NO 616.

XX Human; kidney antigen; immunosuppressive; antiarthritic; antirheumatic;

KW antiproliferative; cyclostatic; cardiant; vasotropic; cerebroprotective;

KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;

KW ophthalmological; antiallergic; hepatotropic; antidiabetic;

KW antiinflammatory; anticancer; vulnery; anticonvulsant; antiparasitic;

KW gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; ds.

XX Homo sapiens.

OS WO20015323-A2.

PN 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US01343.

PF 31-JAN-2000; 2000US-0179065.

PR 04-FEB-2000; 2000US-0180628.

PR 24-FEB-2000; 2000US-0184664.

PR 02-MAR-2000; 2000US-0186350.

PR 16-MAR-2000; 2000US-0188874.

PR 17-MAR-2000; 2000US-0190076.

PR 18-APR-2000; 2000US-0198123.

PR 19-MAY-2000; 2000US-0205515.

PR 07-JUN-2000; 2000US-0209467.

PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.

PR 07-JUL-2000; 2000US-0216647.

PR 11-JUL-2000; 2000US-0216880.

PR 11-JUL-2000; 2000US-0217487.

PR 14-JUL-2000; 2000US-0217496.

PR 26-JUL-2000; 2000US-0218290.

PR 26-JUL-2000; 2000US-0220963.

PR 14-AUG-2000; 2000US-0220964.

PR 14-AUG-2000; 2000US-0224518.

PR 14-AUG-2000; 2000US-0225213.

PR 14-AUG-2000; 2000US-0225214.

PR 14-AUG-2000; 2000US-0225265.

PR 14-AUG-2000; 2000US-0225267.

PR 14-AUG-2000; 2000US-0225268.

PR 14-AUG-2000; 2000US-0225270.

PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.

PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.

PR 18-AUG-2000; 2000US-0226379.

PR 22-AUG-2000; 2000US-0226681.

PR 22-AUG-2000; 2000US-0226868.

PR 22-AUG-2000; 2000US-0227182.

PR 23-AUG-2000; 2000US-0227009.

PR 30-AUG-2000; 2000US-0228924.

PR 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229345.

PR 01-SEP-2000; 2000US-0229345.

PR 05-SEP-2000; 2000US-0229509.

PR 06-SEP-2000; 2000US-0229537.

PR 06-SEP-2000; 2000US-0230438.

PR 08-SEP-2000; 2000US-0231242.

PR 08-SEP-2000; 2000US-0231243.

PR 08-SEP-2000; 2000US-0231244.

PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.

PR 08-SEP-2000; 2000US-0232080.

PR 12-SEP-2000; 2000US-0232081.

PR 12-SEP-2000; 2000US-0231968.

PR 14-SEP-2000; 2000US-0232397.

PR 14-SEP-2000; 2000US-0232398.

PR 14-SEP-2000; 2000US-0232399.

PR 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.

PR 14-SEP-2000; 2000US-0233063.

PR 14-SEP-2000; 2000US-0233064.

PR 21-SEP-2000; 2000US-0233065.

PR 21-SEP-2000; 2000US-0234223.

PR 25-SEP-2000; 2000US-0234274.

PR 25-SEP-2000; 2000US-0234998.

PR 25-SEP-2000; 2000US-0234998.

PR 26-SEP-2000; 2000US-0235834.

PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.

PR 29-SEP-2000; 2000US-0236367.

PR 29-SEP-2000; 2000US-0236368.

PR 29-SEP-2000; 2000US-0236369.

PR 02-OCT-2000; 2000US-0236802.

PR 02-OCT-2000; 2000US-0237037.

PR 02-OCT-2000; 2000US-0237038.

PR 02-OCT-2000; 2000US-0237039.

PR 13-OCT-2000; 2000US-0237040.

PR 13-OCT-2000; 2000US-0239935.

PR 20-OCT-2000; 2000US-0240960.

PR 20-OCT-2000; 2000US-0241221.

PR 20-OCT-2000; 2000US-0241785.

PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.

PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.

PR 08-NOV-2000; 2000US-0244674.

PR 08-NOV-2000; 2000US-0246475.

PR 08-NOV-2000; 2000US-0246476.

PR 08-NOV-2000; 2000US-0246477.

PR 08-NOV-2000; 2000US-0246478.

PR 08-NOV-2000; 2000US-0246523.

PR 08-NOV-2000; 2000US-0246524.

PR 08-NOV-2000; 2000US-0246525.

PR 08-NOV-2000; 2000US-0246526.

PR 08-NOV-2000; 2000US-0246527.

PR 08-NOV-2000; 2000US-0246528.

PR 08-NOV-2000; 2000US-0246532.

PR 08-NOV-2000; 2000US-0246609.

PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.

PR 17-NOV-2000; 2000US-0249207.

PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249254.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 06-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488784/53.

PT New isolated nucleic acids and polypeptides, useful for diagnosing,  
 PT treating and/or preventing human diseases and disorders -

PS Disclosure; SEQ ID NO 616; 564bp + Sequence Listing; English.

CC The invention relates to novel kidney related polynucleotides  
 CC (AA162971-AA163793) and the encoded polypeptides (AAM42417-AAM42691)  
 CC collectively known as kidney antigens and the use of such kidney antigens  
 CC for detecting disorders of the kidney, especially kidney cancer and  
 CC kidney cancer metastases. The polynucleotides and proteins are also  
 CC useful for preventing, treating or ameliorating medical conditions  
 CC e.g. by protein or gene therapy. The genes are isolated from a range  
 CC of human tissues disclosed in the specification. The nucleic acids,  
 CC proteins, antibodies and (ant)agonists are useful in the diagnosis,  
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,  
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,  
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders  
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,  
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple  
 CC sclerosis, rheumatoid arthritis and ulcerative colitis;  
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound  
 CC healing; (e) neurological diseases e.g. cerebral ataxia and epilepsy;  
 CC and (f) infectious diseases such as viral, bacterial, fungal and  
 CC parasitic infections.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 459 BP; 106 A; 108 C; 123 G; 122 T; 0 other;

Alignment Scores:

Pred. No.: 0.424 Length: 459  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 10.31% Indels: 0  
 DB: 22 Gaps: 0

US-09-854-133-586 (1-97) x AA163301 (1-459)

OY 10 SerLeuGIYAspSerGIuFrLeuSerGIu 19

DB 373 AGCCTGGGTGACAGCGAGACCCATCTCA 344

RESULT 38

ABV49991 standard; cDNA; 486 BP.

AC ABV49991.

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 49982.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of  
 PT prostate cells and correlating with presence of prostate cancer, useful  
 PT for detecting presence of prostate cancer, stage of prostate cancer -

PS Claim 1; Page 9741; 11750bp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising  
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the  
 CC specification or its complement. (I) is useful for:  
 CC (a) assessing whether a patient is afflicted with prostate cancer;  
 CC (b) monitoring the progression of prostate cancer in a patient;  
 CC (c) assessing the efficacy of a test compound to inhibit prostate  
 CC cancer in a patient;  
 CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer  
 CC in a patient;  
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;  
 CC (f) assessing the prostate cell carcinogenic potential of a compound;  
 CC (g) determining whether prostate cancer has metastasized in a patient;  
 CC (h) assessing the aggressiveness or indolence of prostate cancer in a  
 CC patient;  
 CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX SQ Sequence 486 BP; 145 A; 114 C; 123 G; 103 T; 1 other;

Alignment Scores:

Pred. No.: 0.448 Length: 486  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0

DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV49991 (1-486)

QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
 DB 444 AGCCTGGGACAGACGAGACTCTCTCA 473

RESULT 39

ABV46023  
 ID ABV46023 standard; cDNA; 490 BP.

AC ABV46023;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker CDNA 46014.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer

PS Claim 1; Page 9092; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 490 BP; 155 A; 116 C; 107 G; 112 T; 0 other;

Alignment Scores:

Pred. No.: 0.451 Length: 490  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV46023 (1-490)

QY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
 DB 283 AGCCTGGGACAGTGAACCCGTCTCA 312

RESULT 40

ABV49206  
 ID ABV49206 standard; cDNA; 491 BP.

AC ABV49206;

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker CDNA 49197.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US05171.

PR 17-FEB-2000; 2000US-183319P.

PR 16-MAR-2000; 2000US-189862P.

PR 25-MAY-2000; 2000US-207454P.

PR 09-JUN-2000; 2000US-211314P.

PR 18-JUL-2000; 2000US-219007P.

PR 13-DEC-2000; 2000US-255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

DR WPI; 2001-662795/76.

PT Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful

PT for detecting presence of prostate cancer, stage of prostate cancer

PS Claim 1; Page 9620; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for:

CC (a) assessing whether a patient is afflicted with prostate cancer;

CC (b) monitoring the progression of prostate cancer in a patient;

CC (c) assessing the efficacy of a test compound to inhibit prostate

CC cancer in a patient;

CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer

CC in a patient;

CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound;

CC (g) determining whether prostate cancer has metastasized in a patient;

CC (h) assessing the aggressiveness or indolence of prostate cancer in a

CC patient;

CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.

XX Sequence 491 BP; 172 A; 82 C; 120 G; 117 T; 0 other;

Alignment Scores:

Pred. No.: 0.452 Length: 491  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 23 Gaps: 0

US-09-854-133-586 (1-97) x ABV49206 (1-491)

OY 10 SerLeuGIYASPSerGIuHrLeuSergIn 19  
|||||

DB 276 AGCTGGCGACGACGAGACCTCTCTCA 305

## RESULT 41

AAS30313/c  
ID AAS30313 standard; cDNA; 497 BP.

XX AAS30313;

XX 21-NOV-2001 (first entry)

XX DNA encoding novel prostate gland antigen, Seq ID No 63.

XX Human; noctropic; neuroprotective; cytostatic; antiparkinsonian;  
XX antianemic; dermatological; immunosuppressive; antiinflammatory;  
XX antihypertensive; antihemorrhagic; virucide; hepatotropic; nephrotropic;  
XX osteopathic; prostate gland; prostatitis; adenocarcinoma; hair loss;  
XX prostatic; carcinoma; prostate neoplastic disorder; urinary system;  
XX hyperplasia; carcinoma; prostate neoplastic disorder; skin aging;  
XX reproductive system disorder; autoimmune disorder; cardiovascular;  
XX systemic lupus erythematosus; rheumatoid arthritis; respiratory;  
XX blood-related disorder; hyperproliferative disorder; respiratory;  
XX neurological disorder; endocrine disorder; inflammatory disorder;  
XX liver disorder; wound healing; food preservative; ss.

XX Homo sapiens.

XX WO200155447-A1.

XX 02-AUG-2001.

XX 17-JAN-2001; 2001MO-US01330.

XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184564.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-APR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225266.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225578.  
XX 14-AUG-2000; 2000US-0225759.  
XX 14-AUG-2000; 2000US-0226279.  
XX 18-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226688.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.

PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0233397.  
PR 14-SEP-2000; 2000US-0233398.  
PR 14-SEP-2000; 2000US-0233399.  
PR 14-SEP-2000; 2000US-0234000.  
PR 14-SEP-2000; 2000US-0234001.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 14-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 21-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 25-SEP-2000; 2000US-0235834.  
PR 26-SEP-2000; 2000US-0235836.  
PR 27-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0244675.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.

PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249219.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251038.  
 PR 05-DEC-2000; 2000US-0251479.  
 PR 06-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 11-DEC-2000; 2000US-0251990.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Barash SC, Ruben SM;  
 WPI: 2001-476223/51.  
 P-PDB; AA018872.

PT Novel isolated prostate gland related polypeptide useful for diagnosis  
 PT and treatment of disorders of prostate such as prostatodystonia,  
 PT prostatosis, prostatitis, benign prostatic hypertrophy and malacoplakia  
 PS Claim 1; SEQ ID No 63; 512pp; English.

CC The invention relates to novel isolated prostate gland related nucleic  
 CC acids (I) and polypeptides (II). (I) and (II) are useful for diagnosis,  
 CC prognosis, prevention, and/or treatment of diseases and/or disorders of  
 CC the prostate such as acute non-bacterial prostatitis, chronic non-  
 CC bacterial prostatitis, acute bacterial prostatitis, prostatodystonia,  
 CC prostatosis, granulomatous prostatitis, malacoplakia, benign prostatic  
 CC adenocarcinomas, transitional cell carcinomas, ductal carcinomas, and  
 CC squamous cell carcinomas. (I), (II) and antibody to (II) are useful for  
 CC diagnosing and treating reproductive system disorders (Paget's disease),  
 CC autoimmune disorders (systemic lupus erythematosus, rheumatoid  
 CC arthritis), blood-related disorders (sickle cell anemia),  
 CC hyperproliferative disorders, urinary system disorders  
 CC (glomerulonephritis), cardiovascular system disorders (arrhythmias),  
 CC respiratory disorders, musculoskeletal system disorders, neural activity  
 CC and neurological disorders (Alzheimer's disease and Parkinson's disease),  
 CC endocrine disorders (Addison's disease), gastrointestinal disorders  
 CC (inflammatory disorders), liver disorders (biliary liver cirrhosis),  
 CC pancreatic and gall bladder disorders, disorders of the large intestine,  
 CC developmental and inherited disorders, diseases at the cellular level,  
 CC and wound healing and epithelial cell proliferation. (I) or (II) is

## Alignment Scores:

Pred. No.: 0.457  
 Score: 10.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 10.31%  
 DB: 22  
 Gaps: 0

US-09-854-133-586 (1-97) x AAS30313 (1-497)

OY 10 SerLeuGlyaspSergLutHLeuSergLn 19  
 Db 184 AGCCTGGTGTGACACTGAGACCCCTGTCTCAA 155  
 RESULT 42  
 ID AAL01994/c  
 ID AAL01994 standard; cDNA; 497 BP.  
 AC AAL01994;  
 XX 21-NOV-2001 (first entry)  
 DT 21-NOV-2001 (first entry)  
 DE Human reproductive system related antigen cDNA SEQ ID NO: 1995.  
 XX Human: reproductive system related antigen; reproductive system disorder;  
 KW cancer; gene therapy; ss.  
 OS Homo sapiens.  
 XX WO20015320-A2.  
 FN 02-AUG-2001.  
 PD 17-JAN-2001; 2001WO-US01339.  
 PF 31-JAN-2000; 2000US-0179065.  
 XX 04-FEB-2000; 2000US-0180628.  
 PR 24-FEB-2000; 2000US-0184664.  
 PR 02-MAR-2000; 2000US-0186350.  
 PR 16-MAR-2000; 2000US-0189874.  
 PR 17-MAR-2000; 2000US-0190076.  
 PR 18-APR-2000; 2000US-0198123.  
 PR 19-MAY-2000; 2000US-0205515.  
 PR 07-JUN-2000; 2000US-0209467.  
 PR 28-JUN-2000; 2000US-0214886.  
 PR 30-JUN-2000; 2000US-0215135.  
 PR 07-JUL-2000; 2000US-0216647.  
 PR 11-JUL-2000; 2000US-0216880.  
 PR 11-JUL-2000; 2000US-0217487.  
 PR 14-JUL-2000; 2000US-0217496.  
 PR 26-JUL-2000; 2000US-0220290.  
 PR 26-JUL-2000; 2000US-0220963.  
 PR 14-AUG-2000; 2000US-0220964.  
 PR 14-AUG-2000; 2000US-0224518.  
 PR 14-AUG-2000; 2000US-0224519.  
 PR 14-AUG-2000; 2000US-0225213.  
 PR 14-AUG-2000; 2000US-0225214.  
 PR 14-AUG-2000; 2000US-0225266.  
 PR 14-AUG-2000; 2000US-0225267.  
 PR 14-AUG-2000; 2000US-0225268.  
 PR 14-AUG-2000; 2000US-0225270.  
 PR 14-AUG-2000; 2000US-0225447.  
 PR 14-AUG-2000; 2000US-0225757.  
 PR 14-AUG-2000; 2000US-0225758.  
 PR 14-AUG-2000; 2000US-0225759.  
 PR 18-AUG-2000; 2000US-0226279.  
 PR 22-AUG-2000; 2000US-0226681.  
 PR 22-AUG-2000; 2000US-0226868.  
 PR 23-AUG-2000; 2000US-0227009.  
 PR 30-AUG-2000; 2000US-0228924.  
 PR 01-SEP-2000; 2000US-0229287.  
 PR 01-SEP-2000; 2000US-0229343.  
 PR 01-SEP-2000; 2000US-0229344.  
 PR 05-SEP-2000; 2000US-0229345.  
 PR 05-SEP-2000; 2000US-0229509.  
 PR 06-SEP-2000; 2000US-0229513.  
 PR 06-SEP-2000; 2000US-0230437.  
 PR 08-SEP-2000; 2000US-0230438.  
 PR 08-SEP-2000; 2000US-0231242.  
 PR 08-SEP-2000; 2000US-0231243.  
 PR 08-SEP-2000; 2000US-0231244.  
 PR 08-SEP-2000; 2000US-0231413.



PR 08-SEP-2000; 2000US-0231414.  
 PR 08-SEP-2000; 2000US-0232080.  
 PR 08-SEP-2000; 2000US-0232081.  
 PR 12-SEP-2000; 2000US-0232397.  
 PR 14-SEP-2000; 2000US-0232398.  
 PR 14-SEP-2000; 2000US-0232399.  
 PR 14-SEP-2000; 2000US-0232400.  
 PR 14-SEP-2000; 2000US-0232401.  
 PR 14-SEP-2000; 2000US-0233063.  
 PR 14-SEP-2000; 2000US-0233064.  
 PR 14-SEP-2000; 2000US-0233065.  
 PR 21-SEP-2000; 2000US-0234274.  
 PR 21-SEP-2000; 2000US-0234275.  
 PR 25-SEP-2000; 2000US-0234998.  
 PR 25-SEP-2000; 2000US-0235484.  
 PR 26-SEP-2000; 2000US-0235834.  
 PR 27-SEP-2000; 2000US-0235836.  
 PR 27-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239335.  
 PR 13-OCT-2000; 2000US-0239337.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246538.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.

PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 PR XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX  
 DR WPI; 2001-465570/50.  
 XX  
 DR P-PSDE; AAM96024.  
 XX  
 PT Isolated nucleic acid molecule encoding a reproductive system antigen  
 XX is used in preventing, treating or ameliorating a medical condition  
 XX  
 PS Claim 1; SEQ ID NO 1995; 1297bp + Sequence Listing; English.  
 XX  
 XX The present invention provides the protein and coding sequences of a  
 CC number of human reproductive system related antigens. These can be used  
 CC in the prevention and treatment of reproductive system disorders,  
 CC including cancer. The present sequence is a coding sequence of the  
 CC invention.  
 XX  
 SQ Sequence 497 BP; 97 A; 136 C; 119 G; 140 T; 5 other;  
 Alignment Scores:  
 Pred. No.: 0.457 Length: 497  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: Gaps: 0  
 US-09-854-133-586 (1-97) x AAL01994 (1-497)  
 QY 10 Serleuglyaspsergluthrleuserglu 19  
 Db 184 AGCTGGGTGACGTGAGACCCCTGTCTCA 155  
 RESULT 43  
 ID AAH09380/C  
 XX AAH09380 standard; cDNA; 509 BP.  
 AC AAH09380;  
 XX  
 DF 26-JUN-2001 (first entry)  
 XX  
 DE Human cDNA clone (3'-primer) SEQ ID NO:6215.  
 XX  
 KW Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN BP1074617-A2.  
 XX  
 PD 07-FEB-2001.  
 XX  
 PF 28-JUL-2000; 2000EP-0116126.  
 XX  
 XX 29-JUL-1999; 99JP-0248036.  
 XX 27-AUG-1999; 99JP-0300253.  
 PR 11-JAN-2000; 2000JP-0118776.

PR 02-MAY-2000; 2000JP-0183767.  
 XX 09-JUN-2000; 2000JP-0241899.  
 PA (HELI-) HELIX RES INST.  
 XX  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 DR WPI: 2001-318749/34.  
 XX  
 PR Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 XX full-length cDNAs.  
 PS  
 PS Claim 3; SEQ ID 6215; 2537bp + CD ROM; English.  
 XX  
 CC The present invention describes primer sets for synthesizing 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of  
 CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC complementary strand of a polynucleotide which comprises a 5'-end  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection and/or diagnosis of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers allow obtaining of the full-length  
 CC cDNAs easily without any specialised methods. AAH0166 to AAH13628 and  
 CC AAH3633 to AAH18742 represent human cDNA sequences; AAH92446 to  
 CC AAH35893 represent human amino acid sequences; and AAH13629 to AAH13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 XX of the present invention.  
 XX  
 SQ Sequence 509 BP; 108 A; 128 C; 131 G; 132 T; 10 other;  
 XX  
 Alignment Scores:  
 Pred. No.:  
 Score: 0.467  
 Percent Similarity: 10.00  
 Best Local Similarity: 100.00%  
 Query Match: 10.00%  
 DB: 10.31%  
 22  
 Gaps: 0  
 0  
 US-09-854-133-586 (1-97) x AAH09380 (1-509)  
 OY 10 SerLeuGlyAspSerGluThrLeuSerGln 19  
 DB 35 AGCTGTGGTGGACAGACGACCTATCTCAA 6  
 RESULT 44  
 ID ABA59875/c  
 XX ABA59875 standard; DNA; 520 BP.  
 AC  
 XX ABA59875;  
 XX  
 DT 01-FEB-2002 (first entry)  
 XX  
 DE Human foetal liver single exon nucleic acid probe #8180.  
 XX  
 KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157277-A2.  
 XX  
 PD 09-AUG-2001.

XX  
 PF 30-JAN-2001; 2001WO-US00669.  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632368.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0236359.  
 PR 04-OCT-2000; 2000US-0024263.  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX WPI: 2001-483447/52.  
 DR  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for  
 PT analysing gene expression in human foetal liver.  
 XX  
 PS  
 PS Claim 1; SEQ ID NO 8180; 639bp + sequence listing; English.  
 XX  
 CC The invention relates to a single exon nucleic acid probe for  
 CC measuring human gene expression in a sample derived from human foetal  
 CC liver. The single exon nucleic acid probes may be used for predicting,  
 CC fetal liver. The present sequence is a single exon nucleic acid  
 CC probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the  
 CC printed specification, but was obtained in electronic format directly  
 CC from WPI at ftp.wipo.int/pub/published\_pcl\_sequences.  
 XX  
 SQ Sequence 520 BP; 135 A; 123 C; 128 G; 134 T; 0 other;  
 XX  
 Alignment Scores:  
 Pred. No.:  
 Score: 0.477  
 Percent Similarity: 10.00  
 Best Local Similarity: 100.00%  
 Query Match: 10.00%  
 DB: 10.31%  
 22  
 Gaps: 0  
 0  
 US-09-854-133-586 (1-97) x ABA59875 (1-520)  
 OY 26 GluArgLysLysLysArgGluArgLysPhe 35  
 DB 262 GAAAGCAAAAAAGAGAGAGAGAAATT 233  
 RESULT 45  
 ID AAK08145/c  
 XX AAK08145 standard; DNA; 520 BP.  
 AC  
 XX AAK08145;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 8136.  
 XX  
 KW Human; brain expressed exon; gene expression analysis; probe;  
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;  
 KW epilepsy; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US00667.  
 PF  
 XX  
 PR 04-FEB-2000; 2000US-0180312.  
 PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.

**Tue May 13 12:12:46 2003**

PR 03-AUG-2000; 2000US-06323366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.  
YY

PI Penn SG, Hanzel DK, Chen W, Rank DR;  
YY

DR WPI; 2001-483446/52.  
XX

PT Single exon nucleic acid probes for analyzing gene expression in human  
PT brains -  
vv

PS Example 4; SEQ ID NO: 8136; 650pp + Sequence Listing; English.  
 YY

CC The present invention provides a number of single exon nucleic acid  
CC probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention.

Sequence 520 BP; 135 A; 123 C; 128 G; 134 T; 0 other;

Alignment Scores:

Pred. No.:	0.477	Length:	520
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	22	Gaps:	0

US-09-854-133-586 (1-97) x AAK08145 (1-520)

26 GluArgLysLysArgGluArgLysPhe 35

Db 262 GAAAGGAAAAAGAGAGAGAGAAATT 233

Search completed: May 11, 2003, 17:16:00  
Job time : 219.735 secs .



GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:16:15 ; Search time 54.0796 Seconds  
(Without alignments)  
550.071 Million cell updates/sec

Title: US-09-854-133-586

Perfect score: 97  
Sequence: 1 EYVSRDMSLSGSESLSQ.....LTGGLCPWATSHLGRKCS 97

Scoring table:  
OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 15338381 residues

Word size: 1

Total number of hits satisfying chosen parameters: 878589

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 60 summaries

Command line parameters:

-MODE=frame+p2n,model -DEV=xlh  
-Q/cgn2\_1/USFTO.spool/US09854133/runat\_05052003\_174133\_728/app\_query.fasta.1.462  
-DB=Issued\_Patents\_NA -OPMT=fastcap -SUFFIX=oligo.rni -MINMATCH=0.1 -DOPEL=0  
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi  
-LIST=60 -DOCALIG=200 -THR.SCORE=quality -THR\_MIN=1 -ALIG=45 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US09854133.ecgn.1.1.47.ernal.05052003\_174133\_728 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAR -LARGEROVER -NEG\_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	9	9.3	36	1	US-08-004-800-18
C 2	9	9.3	36	1	US-08-413-813-18
C 3	9	9.3	36	2	US-08-467-346-18
C 4	9	9.3	1753	4	US-08-205-697A-6
C 5	9	9.3	1753	4	US-08-702-525-6
C 6	9	9.3	1753	4	PCT-US95-02576-6
C 7	9	9.3	4092	4	US-09-306-595C-5
C 8	9	9.3	4136	4	US-09-103-875-2
C 9	9	9.3	12847	1	US-08-550-715-1
C 10	9	9.3	72928	3	US-09-009-913-1
C 11	9	9.3	152331	3	US-09-128-155-16
C 12	8	8.2	36	1	US-08-004-800-16

C 13	8	8.2	36	1	US-08-413-813-16	Sequence 16, Appl
C 14	8	8.2	36	2	US-08-467-346-16	Sequence 16, Appl
C 15	8	8.2	139	4	US-09-084-120-18	Sequence 18, Appl
C 16	8	8.2	291	1	US-07-992-723A-7	Sequence 7, Appl
C 17	8	8.2	291	1	US-07-799-828C-7	Sequence 7, Appl
C 18	8	8.2	291	1	US-08-074-275-7	Sequence 7, Appl
C 19	8	8.2	291	1	US-08-480-366-7	Sequence 7, Appl
C 20	8	8.2	291	2	US-07-952-277A-7	Sequence 7, Appl
C 21	8	8.2	299	2	US-08-623-906A-19	Sequence 19, Appl
C 22	8	8.2	350	2	US-08-623-906A-17	Sequence 17, Appl
C 23	8	8.2	350	3	US-09-157-177-110	Sequence 110, App
C 24	8	8.2	376	2	US-08-623-906A-18	Sequence 18, Appl
C 25	8	8.2	388	2	US-08-623-906A-13	Sequence 18, Appl
C 26	8	8.2	412	4	US-09-018-584A-22	Sequence 22, Appl
C 27	8	8.2	454	2	US-08-623-906A-6	Sequence 6, Appl
C 28	8	8.2	485	4	US-09-385-982-11	Sequence 31, Appl
C 29	8	8.2	484	2	US-08-332-766A-22	Sequence 22, Appl
C 30	8	8.2	1005	3	US-09-103-359-4	Sequence 4, Appl
C 31	8	8.2	1255	1	US-08-097-938-3	Sequence 3, Appl
C 32	8	8.2	1255	1	US-08-476-000-3	Sequence 3, Appl
C 33	8	8.2	1255	1	US-08-472-840-3	Sequence 3, Appl
C 34	8	8.2	1255	2	US-08-476-976-3	Sequence 3, Appl
C 35	8	8.2	1255	3	US-08-474-410-3	Sequence 3, Appl
C 36	8	8.2	1255	4	US-08-486-673B-3	Sequence 3, Appl
C 37	8	8.2	2186	4	US-09-184-001-1	Sequence 1, Appl
C 38	8	8.2	2310	3	US-08-458-434A-1	Sequence 1, Appl
C 39	8	8.2	2415	3	US-09-019-689-1	Sequence 1, Appl
C 40	8	8.2	2558	4	US-09-184-001-3	Sequence 3, Appl
C 41	8	8.2	3892	2	US-08-555-723B-3	Sequence 3, Appl
C 42	8	8.2	3892	3	US-09-123-465-3	Sequence 3, Appl
C 43	8	8.2	4192	4	US-09-122-126B-1	Sequence 1, Appl
C 44	8	8.2	9299	3	US-08-669-161A-29	Sequence 29, Appl
C 45	8	8.2	11461	3	US-09-078-294-7	Sequence 7, Appl
C 46	8	8.2	28720	4	US-09-341-587-7	Sequence 7, Appl
C 47	8	8.2	45546	4	US-09-146-053-6	Sequence 6, Appl
C 48	8	8.2	45546	4	US-09-146-053-6	Sequence 6, Appl
C 49	8	8.2	45546	4	US-09-146-053-6	Sequence 6, Appl
C 50	8	8.2	45716	4	US-08-965-048-5	Sequence 5, Appl
C 51	8	8.2	45965	4	US-08-965-048-5	Sequence 5, Appl
C 52	8	8.2	59065	4	US-09-813-817-3	Sequence 3, Appl
C 53	8	8.2	59065	4	US-09-813-817-3	Sequence 3, Appl
C 54	8	8.2	62804	4	US-09-800-960-3	Sequence 3, Appl
C 55	8	8.2	176372	3	US-09-128-155-17	Sequence 17, Appl
C 56	7	7.2	26	2	US-08-291-011-8	Sequence 8, Appl
C 57	7	7.2	26	4	US-09-266-065-8	Sequence 8, Appl
C 58	7	7.2	36	2	US-08-291-011-4	Sequence 4, Appl
C 59	7	7.2	36	2	US-08-291-011-5	Sequence 5, Appl
C 60	7	7.2	36	4	US-09-266-065-4	Sequence 4, Appl

#### ALIGNMENTS

RESULT 1  
US-08-004-800-18/c  
Sequence 18, Application US/08004800  
Patent No. 5426180  
GENERAL INFORMATION:  
APPLICANT: KOOL, Eric T.  
TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/004,800  
FILING DATE: 19930111  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: McNulty, William E.  
REGISTRATION NUMBER: 22,606  
REFERENCE/DOCKET NUMBER: 80852YX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: circular  
MOLECULE TYPE: DNA  
US-08-004-800-18

Alignment Scores:  
Pred. No.: 0.0797  
Score: 9.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 9.28%  
DB: 1  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-004-800-18 (1-36)  
QY 26 GUARGLYSLYSARGGLUARGLYS 34  
Db 33 GAGAGAAAAAAGAGAGAGAAAAA 7

RESULT 2  
US-08-413-813-18/C  
Sequence 18, Application US/08413813  
Patent No. 5683874  
GENERAL INFORMATION:  
APPLICANT: KOOL, Eric T.  
TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
City: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/413,813  
FILING DATE:  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 80852YX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular

US-08-413-813-18  
Alignment Scores:  
Pred. No.: 0.0797  
Score: 9.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 9.28%  
DB: 1  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-413-813-18 (1-36)  
QY 26 GUARGLYSLYSARGGLUARGLYS 34  
Db 33 GAGAGAAAAAAGAGAGAGAAAAA 7

RESULT 3  
US-08-467-346-18/C  
Sequence 18, Application US/08467346  
Patent No. 5872105  
GENERAL INFORMATION:  
APPLICANT: KOOL, Eric T.  
TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES  
NUMBER OF SEQUENCES: 44  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
City: Garden City  
STATE: New York  
COUNTRY: USA  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,346  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/413,813  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 80852YX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: circular  
US-08-467-346-18

Alignment Scores:  
Pred. No.: 0.0797  
Score: 9.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 9.28%  
DB: 2  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-467-346-18 (1-36)  
QY 26 GUARGLYSLYSARGGLUARGLYS 34  
Db 33 GAGAGAAAAAAGAGAGAGAAAAA 7

RESULT 4  
US-08-205-697A-6/c  
Sequence 6, Application US/08205697A  
Patent No. 6218510  
GENERAL INFORMATION:  
APPLICANT: Sharpe, Arlene H.  
APPLICANT: Borriello, Francescopaulo  
APPLICANT: Freeman, Gordon J.  
APPLICANT: Nadler, Lee M.  
TITLE OF INVENTION: NO. 6218510e1 Forms of T Cell Costimulatory Molecules  
TITLE OF INVENTION: and Uses Therefor  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/205,697A  
FILING DATE: 02-Mar-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: BWI-120  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1753 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-205-697A-6

Alignment Scores:  
Pred. No.: 2.96 Length: 1753  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-205-697A-6 (1-1753)

QY 23 ArgLysLysGluArgLysLysLysArg 31  
|||||  
DB 740 AGAAAAAGAGAGAGAGAGAGAGAGAGAG 714

RESULT 5  
US-08-702-525-6/c  
Sequence 6, Application US/08702525  
Patent No. 6294660  
GENERAL INFORMATION:  
APPLICANT: Sharpe, Arlene H.  
APPLICANT: Borriello, Francescopaulo  
APPLICANT: Freeman, Gordon J.  
APPLICANT: Nadler, Lee M.  
TITLE OF INVENTION: NO. 6294660e1 Forms of T Cell Costimulatory  
TITLE OF INVENTION: Molecules and Uses Therefor  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 State Street  
CITY: Boston

STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/702,525  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/205,697  
FILING DATE: 02-Mar-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207  
REFERENCE/DOCKET NUMBER: BWI-120CPUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617)227-7400  
TELEFAX: (617)227-5941  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1753 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
US-08-702-525-6

Alignment Scores:  
Pred. No.: 2.96 Length: 1753  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-702-525-6 (1-1753)

QY 23 ArgLysLysGluArgLysLysLysArg 31  
|||||  
DB 740 AGAAAAAGAGAGAGAGAGAGAGAGAGAG 714

RESULT 6  
PCT-US95-02576-6/c  
Sequence 6, Application PC/TUS9502576  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Novel Forms of T Cell Costimulatory Molecules  
TITLE OF INVENTION: and Uses Therefor  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 State Street, suite 510  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109-1875  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII Text  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02576  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/205,697  
FILING DATE: 02-Mar-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Mandragouras, Amy E.  
REGISTRATION NUMBER: 36,207

NOTHING

Sequence 1, Application US/08550715



GENERAL INFORMATION:  
APPLICANT: Bowie, Lemuel J.  
TITLE OF INVENTION: Human  $\gamma$ -Thalassemia Mutations as a Predictor of  
TITLE OF INVENTION: Blood-Related Disorders  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS: 13  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 6300 Sears Tower, 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States of America  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/550,715  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Gass, David A.  
REGISTRATION NUMBER: 38,153  
REFERENCE/DOCKET NUMBER: 28493/32834  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/474-6300  
TELEFAX: 312/474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 12847 base pairs  
TYPE: nucleic acid.  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: join(6703..6797, 6915..7119, 7262..7387)  
NAME/KEY: CDS  
LOCATION: join(10514..10608, 10726..10930, 11080..11205)  
US-08-550-715-1  
Alignment Scores:  
Pred. No.: 18.9 Length: 12847  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
Gaps: 0  
DB: 1  
US-09-854-133-586 (1-97) x US-08-550-715-1 (1-12847)  
QY 1 GluValGluValSerArgAspHisala 9  
Db 4510 GAGGTGAGGTGAGCGAGATCAGCGC 4536  
RESULT 10  
US-09-009-913-1/c  
Sequence 1, Application US/09009913  
GENERAL INFORMATION:  
APPLICANT: Axy's Pharmaceuticals, Inc.  
TITLE OF INVENTION: Asthma Related Genes  
NUMBER OF SEQUENCES: 339  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Bozicevic & Reed, LLP  
STREET: 285 Hamilton Ave, Suite 200  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94301

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/009,913  
FILING DATE: 21-JAN-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: SEQ-4P  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-327-3231  
TELEFAX: 650-327-3231  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 72928 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-009-913-1  
Alignment Scores:  
Pred. No.: 94.9 Length: 72928  
Score: 9.00 Matches: 9  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.28% Indels: 0  
Gaps: 0  
DB: 3  
US-09-854-133-586 (1-97) x US-09-009-913-1 (1-72928)  
QY 1 GluValGluValSerArgAspHisala 9  
Db 47805 GAGGTGAGGTGAGCGAGATCAGCGC 47779  
RESULT 11  
US-09-128-155-16  
Sequence 16, Application US/09128155  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
FILE REFERENCE: 09404/052001  
CURRENT APPLICATION NUMBER: US/09/128,155  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 60/091,650  
EARLIER FILING DATE: 1998-07-02  
EARLIER APPLICATION NUMBER: US 60/054,646  
EARLIER FILING DATE: 1997-08-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 152331  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(152331)  
OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16  
Alignment Scores:  
Pred. No.: 188 Length: 152331  
Score: 9.00 Matches: 9

Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 9.28%  
DB: 3  
Conservative: 0  
Mismatches: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-128-155-16 (1-152331)

QY 10 SerLeuGLVAspSerGIuThLeuSer 18  
|||||

Db 118192 AGCCTGGGTGACAGCAGACCTGTCT 118218

RESULT 12

US-08-004-800-16/c

Sequence 16, Application US/08004800

Patent No. 5426180

GENERAL INFORMATION:

APPLICANT: KOOL, Eric T.

TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/004,800

FILING DATE: 19930111

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: McNulty, William E.

REGISTRATION NUMBER: 22,606

REFERENCE/DOCKET NUMBER: 80852Y

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

US-08-004-800-16

Alignment Scores:

Pred. No.: 0.84

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 8.25%

DB: 1

US-09-854-133-586 (1-97) x US-08-004-800-16 (1-36)

QY 27 ArgLYSLysLysArgGLuArgLYs 34  
|||||

Db 35 AGAAAAAGAGAGAGAGAGAAAA 12

RESULT 13

US-08-413-813-16/c

Sequence 16, Application US/08413813

Patent No. 5683874

GENERAL INFORMATION:

APPLICANT: KOOL, Eric T.

TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/413,813

FILING DATE:

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S.

REGISTRATION NUMBER: 31,346

REFERENCE/DOCKET NUMBER: 80852YX

TELECOMMUNICATION INFORMATION:

TELEPHONE: (516) 742-4343

TELEFAX: (516) 742-4366

TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 16:

SEQUENCE CHARACTERISTICS:

LENGTH: 36 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-413-813-16

Alignment Scores:

Pred. No.: 0.84

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 8.25%

DB: 1

US-09-854-133-586 (1-97) x US-08-413-813-16 (1-36)

QY 27 ArgLYSLysLysArgGLuArgLYs 34  
|||||

Db 35 AGAAAAAGAGAGAGAGAGAAAA 12

RESULT 14

US-08-467-346-16/c

Sequence 16, Application US/08467346

Patent No. 5872105

GENERAL INFORMATION:

APPLICANT: KOOL, Eric T.

TITLE OF INVENTION: SINGLE-STRANDED, CIRCULAR OLIGONUCLEOTIDES

NUMBER OF SEQUENCES: 44

CORRESPONDENCE ADDRESS:

ADDRESSEE: Scully, Scott, Murphy & Presser

STREET: 400 Garden City Plaza

CITY: Garden City

STATE: New York

COUNTRY: USA

ZIP: 11530

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,346

FILING DATE: 06-JUN-1995

CLASSIFICATION: 536

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/413,813  
FILING DATE: 30-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: DIGITALO, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 80852YX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-467-346-16

Alignment Scores:  
Pred. No.: 0.84 Length: 36  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: Gaps: 0

US-09-854-133-586 (1-97) x US-08-467-346-16 (1-36)

OY 27 ArgLysLysArgGluArgLys 34  
|||||  
DB 35 AGAAGAGAGAGAGAGAGAGAG 12

RESULT 15  
US-09-084-120-18/c  
Sequence 18, Application US/09084120  
GENERAL INFORMATION:  
APPLICANT: TANG, Jianong  
APPLICANT: MELANCON, Serge B.  
TITLE OF INVENTION: A NOVEL STR MARKER SYSTEM  
TITLE OF INVENTION: FOR DNA FINGERPRINTING  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SMABEY OGILVY RENAULT  
STREET: Suite 1600, 1981 McGill College Avenue  
CITY: Montreal  
STATE: QC  
COUNTRY: Canada  
ZIP: H3A 2Y3  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: Windows  
SOFTWARE: FASTSEQ for Windows Version 2.0b  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/084,120  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: COTE, France  
REGISTRATION NUMBER: 37,037  
REFERENCE/DOCKET NUMBER: 13251-40S FC/1d  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 514 845-7126  
TELEFAX: 514 288-8369  
TELEX:  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 139 base pairs  
TYPE: nucleic acid

STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-09-084-120-18

Alignment Scores:  
Pred. No.: 2.95 Length: 139  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: Gaps: 0

US-09-854-133-586 (1-97) x US-09-084-120-18 (1-139)

OY 23 ArgLysLysGluArgLysLys 30  
|||||  
DB 122 AGAAGAGAGAGAGAGAGAGAG 99

RESULT 16  
US-07-922-723A-7  
Sequence 7, Application US/07922723A  
Patent No. 5369004  
GENERAL INFORMATION:  
APPLICANT: Drs. Michael H. Polymeropoulos  
APPLICANT: and Carl R. Merril  
TITLE OF INVENTION: FIVE HIGHLY INFORMATIVE  
TITLE OF INVENTION: REPEAT POLYMORPHIC DNA MARKERS  
NUMBER OF SEQUENCES: 73  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lowe, Price, LeBlanc & Becker  
STREET: Suite 300, 99 Canal Center Plaza  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: USA  
ZIP: 22314  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: DOS Text File  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/922,723A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: D.J. Mills  
REGISTRATION NUMBER: 34506  
REFERENCE/DOCKET NUMBER: 717081B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703 684 1111  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 291  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-07-922-723A-7

Alignment Scores:  
Pred. No.: 5.87 Length: 291  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: Gaps: 0

US-09-854-133-586 (1-97) x US-07-922-723A-7 (1-291)

OY 23 ArgLysLysGluArgLysLys 30  
|||||  
DB 169 AGAAGAGAGAGAGAGAGAGAG 192

```
RESULT 17
US-07-799-828C-7
; Sequence 7, Application US/07799828C
; Patent No. 5378602
; GENERAL INFORMATION:
; APPLICANT: Drs. Carl R. Merrill and
; APPLICANT: Michael H. Polymeropoulos
; TITLE OF INVENTION: TWENTY SEVEN HIGHLY INFORMATIVE
; TITLE OF INVENTION: MICROSAHELITE REPEAT
; NUMBER OF SEQUENCES: 63
; POLYMORPHIC DNA MARKERS
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, Leblanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: DOS Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/799,828C
; FILING DATE: 19911127
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: D.J. Mullins
; REGISTRATION NUMBER: 34,506
; REFERENCE/DOCKET NUMBER: 717081A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 684 1111
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-07-799-828C-7

Alignment Scores:
Pred. No.: 5.87 Length: 291
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.25% Indels: 0
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-07-799-828C-7 (1-291)
QY 23 ArgLysLysGluArgLysLysLys 30
Db 169 AGAAGAGAAAGAAAGAAAGAAAG 192

RESULT 18
US-08-074-275-7
; Sequence 7, Application US/08074275
; Patent No. 5468610
; GENERAL INFORMATION:
; APPLICANT: Drs. Carl R. Merrill and
; APPLICANT: Michael H. Polymeropoulos
; TITLE OF INVENTION: THREE HIGHLY INFORMATIVE REPEAT
; TITLE OF INVENTION: POLYMORPHIC DNA MARKERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, Leblanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA

RESULT 19
US-08-480-366-7
; Sequence 7, Application US/08480366
; Patent No. 5721100
; GENERAL INFORMATION:
; APPLICANT: Drs. Carl R. Merrill and
; APPLICANT: Michael H. Polymeropoulos
; TITLE OF INVENTION: THREE HIGHLY INFORMATIVE REPEAT
; TITLE OF INVENTION: POLYMORPHIC DNA MARKERS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lowe, Price, Leblanc & Becker
; STREET: Suite 300, 99 Canal Center Plaza
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: DOS Text File
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,366
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: J.G. Mullins
; REGISTRATION NUMBER: 33073
; REFERENCE/DOCKET NUMBER: 717081
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703 684 1111
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 291
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-074-275-7

Alignment Scores:
Pred. No.: 5.87 Length: 291
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.25% Indels: 0
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-074-275-7 (1-291)
QY 23 ArgLysLysGluArgLysLysLys 30
Db 169 AGAAGAGAAAGAAAGAAAGAAAG 192
```

TELEPHONE: 703 684 1111  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 291  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-480-366-7

## Alignment Scores:

Pred. No.:	5.87	Length:	291
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.25%	Indels:	0
DB:	1	Gaps:	0

US-09-854-133-586 (1-97) x US-08-480-366-7 (1-291)

OY 23 ArgLysLysGluArgLysLysLys 30

DB 169 AGAAGAGAAAGAGAGAGAGAGAG 192

## RESULT 20

US-07-952-277A-7  
Sequence 7, Application US/07952277A

Patent No. 5861504

GENERAL INFORMATION:

APPLICANT: Drs. Michael H. Polymeropoulos

APPLICANT: and Carl R. Merrill

TITLE OF INVENTION: ELEVEN HIGHLY INFORMATIVE

NUMBER OF SEQUENCES: 85

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lowe, Price, Leblanc &amp; Becker

STREET: Suite 300, 99 Canal Center Plaza

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: DOS Text File

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/952,277A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: D.J. Mills

REGISTRATION NUMBER: 34506

REFERENCE/DOCKET NUMBER: 717081C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703 684 1111

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 291

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

US-07-952-277A-7

## Alignment Scores:

Pred. No.:	5.87	Length:	291
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.25%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x US-07-952-277A-7 (1-291)

OY 23 ArgLysLysGluArgLysLysLys 30

DB 169 AGAAGAGAAAGAGAGAGAGAGAG 192

## RESULT 21

US-08-623-906A-19

Sequence 19, Application US/08623906A

Patent No. 5874217

GENERAL INFORMATION:

APPLICANT: Stevenson, Tamara

APPLICANT: Dvorak, Jan

TITLE OF INVENTION: Microsatellite Sequences for Canine

NUMBER OF SEQUENCES: 60

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON &amp; HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: CA

COUNTRY: US

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/623,906A

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sherwood, Pamela J.

REGISTRATION NUMBER: 36,677

REFERENCE/DOCKET NUMBER: A-62282/BIR

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989

TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 19:

SEQUENCE CHARACTERISTICS:

LENGTH: 299 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1..128

OTHER INFORMATION: /note="Nucleotides 1-128 are

OTHER INFORMATION: unique flanking sequence"

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 129..199

OTHER INFORMATION: /note="Nucleotides 129-199 are

OTHER INFORMATION: repeat sequence"

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 200..299

OTHER INFORMATION: /note="Nucleotides 200-299 are

OTHER INFORMATION: unique flanking sequence"

US-08-623-906A-19

## Alignment Scores:

Pred. No.:	6.02	Length:	299
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.25%	Indels:	0
DB:	2	Gaps:	0

US-09-854-133-586 (1-97) x US-08-623-906A-19 (1-299)

OY 23 Arglyslgsluarglyslgsls 30  
|||||  
Db 176 AGAAGAGAGAGAGAGAGAGAGAG 199

## RESULT 22

US-08-623-906A-17  
; Sequence 17, Application US/08623906A  
; Patent No. 5874217

## GENERAL INFORMATION:

APPLICANT: Stevenson, Tamara  
APPLICANT: Dvorak, Jan  
APPLICANT: Halverson, Joy  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/623,906A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J.  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: A-62282/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249

INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 1..109  
OTHER INFORMATION: /note= "Nucleotides 1-109 are  
OTHER INFORMATION: unique flanking sequence"  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 110..195  
OTHER INFORMATION: /note= "Nucleotides 110-195 are  
OTHER INFORMATION: repeat sequence"  
FEATURE:  
NAME/KEY: misc.feature  
LOCATION: 196..350  
OTHER INFORMATION: /note= "Nucleotides 196-350 are  
OTHER INFORMATION: unique flanking sequence"  
US-08-623-906A-17

## Alignment Scores:

Pred. No.: 6.96 Length: 350  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-623-906A-17 (1-350)

OY 23 Arglyslgsluarglyslgsls 30  
|||||  
Db 118 AGAAGAGAGAGAGAGAGAGAGAG 141

## RESULT 23

US-09-157-177-110  
; Sequence 110, Application US/09157177  
; Patent No. 6090558

## GENERAL INFORMATION:

APPLICANT: Butler, John M.  
APPLICANT: Li, Jia  
APPLICANT: Monforte, Joseph A.  
APPLICANT: Becker, Christopher H.  
TITLE OF INVENTION: DNA TYPING BY MASS SPECTROMETRY WITH POLYMORPHIC DNA  
FILE REFERENCE: GENE/017/GETR017P  
CURRENT APPLICATION NUMBER: US/09/157,177  
CURRENT FILING DATE: 1998-09-18  
EARLIER APPLICATION NUMBER: 60/059,415  
EARLIER FILING DATE: 1997-09-19  
NUMBER OF SEQ ID NOS: 135  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 110  
LENGTH: 350  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-157-177-110

## Alignment Scores:

Pred. No.: 6.96 Length: 350  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 3 Gaps: 0

US-09-854-133-586 (1-97) x US-09-157-177-110 (1-350)

OY 22 leuArglyslgsluarglyslgsls 29  
|||||  
Db 70 CTCGAGAGAGAGAGAGAGAGAGAG 93

## RESULT 24

US-08-623-906A-18  
; Sequence 18, Application US/08623906A  
; Patent No. 5874217

## GENERAL INFORMATION:

APPLICANT: Stevenson, Tamara  
APPLICANT: Dvorak, Jan  
APPLICANT: Halverson, Joy  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/623,906A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J.  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: A-62282/BIR

US-09-854-133-586 (1-97) x US-08-623-906A-17 (1-350)

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 376 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..79  
OTHER INFORMATION: /note= "Nucleotides 1-79 are unique  
OTHER INFORMATION: flanking sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 80..229  
OTHER INFORMATION: /note= "Nucleotides 80-229 are  
OTHER INFORMATION: repeat sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 230..376  
OTHER INFORMATION: /note= "Nucleotides 230-376 are  
OTHER INFORMATION: unique flanking sequence"  
US-08-623-906A-18  
Alignment Scores:  
Pred. No.: 7.45 Length: 376  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0  
US-09-854-133-586 (1-97) x US-08-623-906A-18 (1-376)  
QY 23 ArgLysLysGluArgLysLysLys 30  
Db 98 AGAAGAGAAAGAAAGAGAGAGAA 121  
RESULT 25  
US-08-623-906A-13/C  
Sequence 13, Application US/08623906A  
Patent No. 5874217  
GENERAL INFORMATION:  
APPLICANT: Stevenson, Tamara  
APPLICANT: Dvorak, Jan  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FIEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/623,906A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J.  
REGISTRATION NUMBER: 36,677  
REFERENCE/DOCKET NUMBER: A-62282/B1R  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 388 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..131  
OTHER INFORMATION: /note= "Nucleotides 1-131 are  
OTHER INFORMATION: unique flanking sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 132..235  
OTHER INFORMATION: /note= "Nucleotides 132-235 are  
OTHER INFORMATION: repeat sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 236..388  
OTHER INFORMATION: /note= "Nucleotides 236-388 are  
OTHER INFORMATION: unique flanking sequence"  
US-08-623-906A-13  
Alignment Scores:  
Pred. No.: 7.67 Length: 388  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0  
US-09-854-133-586 (1-97) x US-08-623-906A-13 (1-388)  
QY 23 ArgLysLysGluArgLysLysLys 30  
Db 176 AGAAGAGAAAGAAAGAGAGAGAA 153  
RESULT 26  
US-09-018-584A-22  
Sequence 22, Application US/09018584A  
Patent No. 6238863  
GENERAL INFORMATION:  
APPLICANT: Schumm, James W.  
APPLICANT: Bacher, Jeffrey W.  
TITLE OF INVENTION: MATERIALS AND METHODS FOR  
TITLE OF INVENTION: IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM  
NUMBER OF SEQUENCES: 147  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Promega Corporation  
STREET: 2800 Woods Hollow Road  
CITY: Madison  
STATE: Wisconsin  
COUNTRY: U.S.A.  
ZIP: 53711-5399  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 MB  
COMPUTER: IBM compatible PC  
OPERATING SYSTEM: Windows 95  
SOFTWARE: Word 97 (DOS text format)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/018,584A  
FILING DATE: 04-Feb-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Grady J. Frenchick  
REGISTRATION NUMBER: 29,018  
REFERENCE/DOCKET NUMBER: 16026.9180  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (608) 257-3501

TELEFAX: (608) 257-2275  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 412 bp  
TYPE: Nucleic Acid  
STRANDEDNESS: Double  
TOPOLOGY: Circular  
MOLECULE TYPE: Genomic DNA  
HYPOTHETICAL: no  
IMMEDIATE SOURCE:  
LIBRARY: plasmid, pGem3zf(+)  
CLONE: G234  
POSITION IN GENOME:  
CHROMOSOME/SEGMENT: 16 qter  
US-09-018-584A-22

Alignment Scores:  
Pred. No.: 8.11 Length: 412  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-09-018-584A-22 (1-412)

OY 22 leuarglysgluarglylslys 29  
DB 49 CTCAGAGAAAAAGAAAGAAAAA 72

RESULT 27  
US-08-623-906A-6  
Sequence 6, Application US/08623906A  
Patent No. 5874217  
GENERAL INFORMATION:  
APPLICANT: Stevenson, Tamara  
APPLICANT: Dvorak, Jan  
TITLE OF INVENTION: Microsatellite Sequences for Canine  
TITLE OF INVENTION: Genotyping  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT  
STREET: 4 Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/623,906A  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Sherwood, Pamela J.  
REGISTRATION NUMBER: 36,677  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 454 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: misc\_feature

LOCATION: 1..133  
OTHER INFORMATION: /note="Nucleotides 1-133 are  
OTHER INFORMATION: unique flanking sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 134..307  
OTHER INFORMATION: /note="Nucleotides 134-207 are  
OTHER INFORMATION: repeat sequence"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 308..454  
OTHER INFORMATION: /note="Nucleotides 308-454 are  
OTHER INFORMATION: unique flanking sequence"  
US-08-623-906A-6

Alignment Scores:  
Pred. No.: 8.87 Length: 454  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 2 Gaps: 0

US-09-854-133-586 (1-97) x US-08-623-906A-6 (1-454)

OY 23 Arglyslsgluarglylslys 30  
DB 216 AGAAGAGAAAGAAAGAAAGAAAG 239

RESULT 28  
US-09-385-982-31  
Sequence 31, Application US/09385982  
Patent No. 6262334  
GENERAL INFORMATION:  
APPLICANT: ENDEGE, WILSON O., ET AL.  
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION  
FILE REFERENCE: CCDDNA-260XX  
CURRENT APPLICATION NUMBER: US/09/385,982  
CURRENT FILING DATE: 1999-08-30  
EARLIER APPLICATION NUMBER: 09/328,111  
EARLIER FILING DATE: 1999-06-08  
EARLIER APPLICATION NUMBER: 60/117,393  
EARLIER FILING DATE: 1999-01-27  
EARLIER APPLICATION NUMBER: 60/098,639  
EARLIER FILING DATE: 1998-08-31  
NUMBER OF SEQ ID NOS: 544  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 31  
LENGTH: 485  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(485)  
OTHER INFORMATION: n = A,T,C or G  
US-09-385-982-31

Alignment Scores:  
Pred. No.: 9.43 Length: 485  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-09-385-982-31 (1-485)

OY 23 Arglyslsgluarglylslys 30  
DB 136 AGAAGAGAAAGAAAGAAAGAAAG 159

RESULT 29



```

US-08-332-766A-22
; Sequence 22, Application US/08332766A
; Patent No. 5843647
GENERAL INFORMATION:
APPLICANT: JEFFREYS, Alec J.
APPLICANT: ARMOUR, John
TITLE OF INVENTION: SIMPLE TANDEM REPEATS
NUMBER OF SEQUENCES: 125
ADDRESSSEE: CUSHMAN DABRY & CUSHMAN, L.L.P.
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/332,766A
FILING DATE: 01-NOV-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9326052.9
FILING DATE: 21-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: BIRD, Donald J.
REGISTRATION NUMBER: 25,323
REFERENCE/DOCKET NUMBER: 217211/M94/0434/GB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3000
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-332-766A-22
Alignment Scores:
Pred. No.: 9.6 Length: 494
Score: 8.0 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.25% Indels: 0
Gaps: 0
DB: 2
US-09-854-133-586 (1-97) x US-08-332-766A-22 (1-494)
QY 23 ArgLysLysGUArgLysLysLys 30
|||||
Db 108 AGAAGAAGAAAGAAAGAAAGAA 131
RESULT 30
US-09-103-359-4
; Sequence 4, Application US/09103359
; Patent No. 6057108
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Guegler, Karl J.
APPLICANT: Corley, Neil C.
APPLICANT: Yue, Henry
APPLICANT: Patterson, Chandra
TITLE OF INVENTION: HUMAN ARF-RELATED PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
```

```

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
FILING DATE: HEREWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Cerrone, Michael C
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0537 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-855-0572
TELEX:

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1005 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: COLNNOT13
CLONE: 1333754
US-09-103-359-4

Alignment Scores:
Pred. No.:      18.6      Length:      1005
Score:          8.00      Matches:      8
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match:      8.25%   Indels:      0
DB:               3      Gaps:       0

US-09-854-133-586 (1-97) x US-09-103-359-4 (1-1005)

Oy      23 ArglyslYsgUAWrlyslYslYslYs 30
      |||||||
Db      115 AGAAGAAAGAAAGAAAGAAAGAAA 138

RESULT 31
US-08-097-938-3
Sequence 3, Application US/08097938
Patent No. 5629174

GENERAL INFORMATION:
APPLICANT: SUNDELIN, JOHAN
APPLICANT: SCARBOROUGH, ROBERT M.
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR AND ITS
TITLE OF INVENTION: AGONISTS AND ANTAGONISTS
NUMBER OF SEQUENCES: 59
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20006-1812

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/097,938

```

FILING DATE: 26-JUL-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURASHIGE, KATE H.  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 22803-20006.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-097-938-3

Alignment Scores:  
Pred. No.: 22.8 Length: 1255  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-097-938-3 (1-1255)

QY 11 LeuGLYAspSerGluThrIeuSer 18  
|||||  
Db 11 CTGGGTGACGAGAGACCTGTCT 34

RESULT 32  
US-08-476-000-3  
Sequence 3, Application US/08476000  
Patent No. 5716789  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,000  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-476-000-3

Alignment Scores:  
Pred. No.: 22.8 Length: 1255  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 1 Gaps: 0

US-09-854-133-586 (1-97) x US-08-476-000-3 (1-1255)

QY 11 LeuGLYAspSerGluThrIeuSer 18  
|||||  
Db 11 CTGGGTGACGAGAGACCTGTCT 34

RESULT 33  
US-08-472-840-3  
Sequence 3, Application US/08472840  
Patent No. 5763575  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,840  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-472-840-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 1  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-472-840-3 (1-1255)

OY 11 LeuGlyAspSerGluThrLeuSer 18  
DB 11 CTGGGTGACGACGAGACCTGTCT 34

RESULT 34  
US-08-476-976-3  
Sequence 3, Application US/08476976  
Patent No. 5874400  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,976  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-476-976-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 2  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-476-976-3 (1-1255)

OY 11 LeuGlyAspSerGluThrLeuSer 18  
DB 11 CTGGGTGACGACGAGACCTGTCT 34

RESULT 35  
US-08-474-410-3  
Sequence 3, Application US/08474410  
Patent No. 6043212  
GENERAL INFORMATION:  
APPLICANT: SUNDELIN, JOHAN  
APPLICANT: SCARBOROUGH, ROBERT M.  
TITLE OF INVENTION: RECOMBINANT C140 RECEPTOR, ITS AGONISTS  
TITLE OF INVENTION: AND ANTAGONISTS, AND NUCLEIC ACIDS ENCODING THE RECEPTOR  
NUMBER OF SEQUENCES: 63  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Ave. N.W., Ste. 5500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20006-1812  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/474,410  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/390,301  
FILING DATE: 25-JAN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: ADLER, REID G.  
REGISTRATION NUMBER: 30,988  
REFERENCE/DOCKET NUMBER: 2803-0006.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 887-1500  
TELEFAX: (202) 887-0763  
TELEX: 90-4030  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1255 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 56..1249  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 56  
US-08-474-410-3

Alignment Scores:  
Pred. No.: 22.8  
Score: 8.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 8.25%  
DB: 3  
Length: 1255  
Matches: 8  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-08-474-410-3 (1-1255)

OY 11 LeuGlyAspSerGluThrLeuSer 18  
DB 11 CTGGGTGACAGCGAGACCTGTCT 34

RESULT 36

US-08-486-673B-3  
; Sequence 3, Application US/08486673B  
; Patent No. 6297026  
; GENERAL INFORMATION:

APPLICANT: Sundelin, Johan  
APPLICANT: Scarborough, Robert M.

TITLE OF INVENTION: Nucleic Acids Encoding the C140 Receptor  
FILE REFERENCE: 44481-5006-08-US  
CURRENT APPLICATION NUMBER: US/08/486,673B

PRIOR FILING DATE: 1995-06-07  
PRIOR APPLICATION NUMBER: US 08/097,938

PRIOR FILING DATE: 1993-07-26  
PRIOR APPLICATION NUMBER: PCT/US94/08536

NUMBER OF SEQ ID NOS: 63  
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 3  
LENGTH: 1255

TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: CDS  
LOCATION: (56)..(1249)

OTHER INFORMATION: C140 receptor, genomic DNA and deduced protein  
US-08-486-673B-3

Alignment Scores:

Pred. No.:	22.8	Length:	1255
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.25%	Indels:	0
DB:	4	Gaps:	0

US-09-854-133-586 (1-97) x US-08-486-673B-3 (1-1255)

OY 11 LeuGlyAspSerGluThrLeuSer 18  
DB 11 CTGGGTGACAGCGAGACCTGTCT 34

RESULT 37

US-09-184-001-1/c  
; Sequence 1, Application US/09184001A  
; Patent No. 6303333  
; GENERAL INFORMATION:

APPLICANT: MICHALOVICH, DAVID  
APPLICANT: SIMS, MATTHEW A.

TITLE OF INVENTION: NOVEL COMPOUNDS  
FILE REFERENCE: GP-30088

CURRENT APPLICATION NUMBER: US/09/184,001A  
EARLIER FILING DATE: 1998-11-02

EARLIER APPLICATION NUMBER: 9817479.0  
EARLIER FILING DATE: 1998-08-11

EARLIER APPLICATION NUMBER: 9806221.9  
NUMBER OF SEQ ID NOS: 4

SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1

LENGTH: 2186  
TYPE: DNA

ORGANISM: HOMO SAPIENS  
US-09-184-001-1

Alignment Scores:

Pred. No.:	38.3	Length:	2186
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.25%	Indels:	0
DB:	4	Gaps:	0

US-09-854-133-586 (1-97) x US-09-184-001-1 (1-2186)

OY 87 ThrArgSerHisLeuGlyArg 94  
DB 215 ACCAGAGCCACTGGGGAAG 192

RESULT 38

US-08-458-434A-1  
; Sequence 1, Application US/08458434A  
; Patent No. 6083690  
; GENERAL INFORMATION:

APPLICANT: Harris Ph.D., Stephen E.  
APPLICANT: Mundy M.D., Gregory R.

APPLICANT: Gosh-Choudhury Ph.D., Nandini  
APPLICANT: Feng Ph.D., Jian Q.

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:

ADDRESSEE: James C. Weseman, Esq.  
STREET: 401 B. Street, Suite 1700

CITY: San Diego  
STATE: CA

COUNTRY: USA  
ZIP: 92101

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/458,434A  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Weseman, James C.

REGISTRATION NUMBER: 30,507  
REFERENCE/DOCKET NUMBER: P00060US0

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 699-3604

TELEFAX: 619-236-1048  
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
LENGTH: 2310 base pairs  
TYPE: nucleic acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)

FEATURE:  
NAME/KEY: CDS  
LOCATION: 768..1991

US-08-458-434A-1

Alignment Scores:

Pred. No.:	40.3	Length:	2310
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	8.25%	Indels:	0
DB:	3	Gaps:	0

US-09-854-133-586 (1-97) x US-08-458-434A-1 (1-2310)

OY 23 ArgLysLysGluArgLysLys 30  
DB 2143 AGAAGAGAAAGAAAGAAAG 2166

RESULT 39  
US-09-019-689-1/c  
Sequence 1, Application US/09019689  
Patent No. 6103527  
GENERAL INFORMATION:  
APPLICANT: GRAULICH, wolff  
APPLICANT: NETTELBECK, Dirk  
APPLICANT: SEDLACER, Hans-Harald  
APPLICANT: MUELLER, Rolf  
TITLE OF INVENTION: PROMOTER OF THE HUMAN ENDOGLIN GENE AND  
TITLE OF INVENTION: ITS USE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY & LARDNER  
STREET: 3000 K Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/019,689  
FILING DATE: 06-FEB-1998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 19704301.1  
FILING DATE: 06-FEB-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Sandercocck, Colin G.  
REGISTRATION NUMBER: 31,298  
REFERENCE/DOCKET NUMBER: 016779/0130  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 672-5300  
TELEFAX: (202) 672-5399  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2415 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-09-019-689-1

Alignment Scores:  
Pred. No.: 42 Length: 2415  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 3 Gaps: 0

US-09-854-133-586 (1-97) x US-09-019-689-1 (1-2415)  
QY 24 LysLysGluArgLysLysArg 31  
DB 1381 AAAAAAAAAAAAAAAAAAGA 1358

RESULT 40  
US-09-184-001-3/c  
Sequence 3, Application US/09184001A  
Patent No. 6303333  
GENERAL INFORMATION:  
APPLICANT: MICHALOVICH, DAVID  
APPLICANT: SIMS, MATTHEW A.  
APPLICANT: SHAIKH, NARJIS  
TITLE OF INVENTION: NOVEL COMPOUNDS  
FILE REFERENCE: GP-30088  
CURRENT APPLICATION NUMBER: US/09/184,001A

CURRENT FILING DATE: 1998-11-02  
EARLIER APPLICATION NUMBER: 9817479.0  
EARLIER FILING DATE: 1998-08-11  
EARLIER APPLICATION NUMBER: 9806221.9  
EARLIER FILING DATE: 1998-03-23  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 2558  
TYPE: DNA  
ORGANISM: HOMO SAPIENS  
FEATURE:  
NAME/KEY: UNSURE  
LOCATION: (146)(161)(178)(233)(254)(296)  
US-09-184-001-3

Alignment Scores:  
Pred. No.: 44.3 Length: 2558  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 8.25% Indels: 0  
DB: 4 Gaps: 0

US-09-854-133-586 (1-97) x US-09-184-001-3 (1-2558)  
QY 87 ThrArgSerHisLeuGlyArg 94  
DB 111 ACCAGAGCAGCATCTGGGAGAGG 88

RESULT 41  
US-08-555-723B-3/c  
Sequence 3, Application US/08555723B  
Patent No. 5837534  
GENERAL INFORMATION:  
APPLICANT: OLSON, Eric N.  
APPLICANT: LI, Li  
APPLICANT: MIANO, Joseph M.  
TITLE OF INVENTION: SMOOTH MUSCLE 22 PROMOTER, GENE TRANSFER  
TITLE OF INVENTION: VECTORS CONTAINING THE SAME, AND METHOD OF  
TITLE OF INVENTION: USE OF THE SAME TO TARGET GENE EXPRESSION IN  
TITLE OF INVENTION: ARTERIAL SMOOTH MUSCLE CELLS  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SUGHRUE, MIOM, ZINN, MACPEAK & SEAS  
STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
CITY: Washington, D.C.  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/555,723B  
FILING DATE: 14-NOVEMBER-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: KIT, Gordon  
REGISTRATION NUMBER: 30,764  
REFERENCE/DOCKET NUMBER: A-6663  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 293-7060  
TELEFAX: (202) 293-7860  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3892 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: genomic

```
DB:                               3                                Gaps: 0

US-09-854-133-586 (1-97) x US-09-123-465-3 (1-3892)

QY      27  ArgLysLysArgGluArgLys 34
         |||||
Db       3448 AGAAGAAAAAGAAAGAAAAGAAA 3425

RESULT 43
US-09-122-126B-1/c
; Sequence 1, Application US/09122126B
; Patent No. 6451575
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: AGGRECAN DEGRADING METALLO PROTEASES
; FILE REFERENCE: DM6909
; CURRENT APPLICATION NUMBER: US/09/122,126B
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 4192
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (406)..(2916)
US-09-122-126B-1

Alignment Scores:
Pred. No.:          70.1           Length:        4192
Percent:            8.00           Matches:         8
Score Similarity:   100.00%        Conservative:    0
Best Local Similarity: 100.00%     Mismatches:     0
Query Match:        8.25%          Indels:         0
DB:                  4              Gaps:           0

US-09-854-133-586 (1-97) x US-09-122-126B-1 (1-4192)

QY      23  ArgLysLysGluArgLysLys 30
         |||||
Db       3533 AGAAGAAAGAAAGAAAAA 3510

RESULT 44
US-08-458-434A-7
; Sequence 7, Application US/08458434A
; Patent No. 6083690
; GENERAL INFORMATION:
; APPLICANT: Harris Ph.D., Stephen E.
; APPLICANT: Mundy M.D., Gregory R.
; APPLICANT: Gosh-Choudhury Ph.D., Nandini
; APPLICANT: Feng Ph.D., Jian Q.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR IDENTIFYING
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: James C. Weseman, Esq.
; STREET: 401 B. Street, Suite I700
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,434A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Weseman, James C.
```

```

? TELEFAX: (212) 838-3884
? INFORMATION FOR SEQ ID NO: 29:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 11461 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
US-08-6659-161A-29

Alignment Scores:
Pred. No.: 179
Score: 8.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 8.25%
DB: 3

US-09-8554-133-586 (1-97) x US-08-6659-161A-29 (1-11461)

      Length: 11461
      Matches: 8
      Conservative: 0
      Mismatches: 0
      Indels: 0
      Gaps: 0

```

Search completed: May 11, 2003, 17:17:42  
Job time : 87.0797 secs

```
QY 24 LSLYSGIUA RGLYSLYSLYSARG 31
      |||||
Db 10461 AAGAAAGAAAGAAAAAAGAAAGA 10438
```





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 17:18:00 ; Search time 115.027 Seconds  
(without alignments)  
1047.953 Million cell updates/sec

Title: US-09-854-133-586  
Sequence: 1 EVEVSRDHSALGSESLISQT.....ITGGLCWATSHLRKKS 97

Scoring table:  
OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 783854 seqs, 621352466 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1565805

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 60 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO.spool/US09854133/runat\_05052003\_174135\_809/app.query.fasta.1.462  
-DB=Published.Applications\_NA -OFMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo  
-TRANS=human40.cdt -LIST=60 -DOCALIGN=200 -THR.SCORE=quality -THR.MIN=1  
-ALIGN=45 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09854133 -ECGN\_1.1.117 -runat\_05052003\_174135\_809  
-NCPU=6 -ICPU=3 -NO\_XLPEX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:

- 1: /cgn2\_6/ptodata/2/pubna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubna/US09\_NEW\_PUB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubna/US10\_NEW\_PUB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubna/US10\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubna/US60\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	97	100.0	337 9 US-09-854-133-442	Sequence 442, App
2	97	100.0	337 10 US-09-738-973-442	Sequence 442, App
3	96	99.0	5981 9 US-09-854-133-441	Sequence 441, App
4	96	99.0	5981 10 US-09-738-973-441	Sequence 441, App

5	93	95.9	2239 9 US-09-854-133-440	Sequence 440, App
6	93	95.9	2239 10 US-09-738-973-440	Sequence 440, App
7	93	95.9	520 9 US-10-163-866-32	Sequence 32, App
8	37	38.1	572 9 US-10-046-935-1307	Sequence 1307, App
9	37	38.1	572 9 US-09-878-178-1307	Sequence 1307, App
10	37	38.1	572 9 US-10-146-502-1307	Sequence 52, App
11	37	38.1	1268 9 US-10-163-866-53	Sequence 53, App
12	37	38.1	1528 9 US-10-163-866-52	Sequence 52, App
13	37	38.1	1542 9 US-10-163-866-33	Sequence 33, App
14	37	38.1	1861 9 US-10-163-866-30	Sequence 30, App
15	37	38.1	2000 9 US-10-163-866-24	Sequence 29, App
16	37	38.1	2482 9 US-10-163-866-29	Sequence 31, App
17	37	38.1	3144 9 US-09-764-897-3626	Sequence 326, App
18	12	12.4	29607 10 US-09-933-797-216	Sequence 216, App
19	11	11.3	294 9 US-10-091-504-2209	Sequence 2209, App
20	11	11.3	32190 9 US-09-764-869-2209	Sequence 2201, App
21	11	11.3	32190 10 US-09-998-598-2201	Sequence 71, App
22	10	10.3	261 9 US-09-803-719-71	Sequence 71, App
23	10	10.3	291 9 US-09-803-719-334	Sequence 234, App
24	10	10.3	291 9 US-09-803-719-83	Sequence 83, App
25	10	10.3	303 9 US-09-954-531-403	Sequence 403, App
26	10	10.3	351 9 US-09-918-995-12641	Sequence 12641, App
27	10	10.3	418 10 US-09-918-995-12641	Sequence 12641, App
28	10	10.3	483 9 US-09-764-891-1995	Sequence 1995, App
29	10	10.3	497 9 US-09-864-761-12019	Sequence 12019, App
30	10	10.3	520 10 US-09-918-995-13252	Sequence 13252, App
31	10	10.3	554 9 US-10-028-072-255	Sequence 255, App
32	10	10.3	1432 9 US-10-121-049-255	Sequence 255, App
33	10	10.3	1432 9 US-10-123-904-255	Sequence 255, App
34	10	10.3	1432 9 US-10-140-070-255	Sequence 255, App
35	10	10.3	1432 9 US-10-176-921-255	Sequence 255, App
36	10	10.3	1432 9 US-10-176-921-255	Sequence 255, App
37	10	10.3	1432 9 US-10-176-921-255	Sequence 255, App
38	10	10.3	1432 9 US-10-176-921-255	Sequence 255, App
39	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
40	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
41	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
42	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
43	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
44	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
45	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
46	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
47	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
48	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
49	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
50	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
51	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
52	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
53	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
54	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
55	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
56	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
57	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
58	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
59	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App
60	10	10.3	1432 9 US-10-140-474-255	Sequence 255, App

## ALIGNMENTS

RESULT 1  
US-09-854-133-442  
Sequence 442, Application US/09854133  
Publication No. US20020183499A1  
GENERAL INFORMATION:  
APPLICANT: Lodes, Michael J.  
APPLICANT: Mohamath, Raodoh  
APPLICANT: Henderson, Robert A.  
APPLICANT: Benson, Darin R.  
APPLICANT: Secrist, Heather  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
THE THERAPY AND DIAGNOSIS OF LUNG CANCER  
FILE REFERENCE: 210121.475C10

; CURRENT APPLICATION NUMBER: US/09/854,133  
 ; CURRENT FILING DATE: 2001-05-11  
 ; NUMBER OF SEQ ID NOS: 735  
 ; SOFTWARE: FASTSEQ for Windows Version 3.0  
 ; SEQ ID NO 442  
 ; LENGTH: 337  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 US-09-854-133-442

## Alignment Scores:

Pred. No.: 3,54e-94 Length: 337  
 Score: 97.00 Matches: 97  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-854-133-586 (1-97) x US-09-854-133-442 (1-337)

QY 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
 Db 5 GAGTTTAAAGTGAAGACAGATCATGCGCTGGGTGACAGTGAAGTCTGTCTCAACA 64  
 QY 21 GluLeuArgLysGlyGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 40  
 Db 65 GAATTAAGGAAAAAGAAAAAGAGAGAGAGAAATTCAGGCCAATGTGTGCG 124  
 QY 41 IleAspPheIlePheTrpIlePheTrpIleLeuLeuPheSerHisIleTrpIleGln 60  
 Db 125 ATGATTTTATCATATTCGTGATTTTGTGATTTCTTTTCTCATCATCGATTGCG 184  
 QY 61 GluSerLeuLeuGlyProPheSerProLysGluValThrCysArgGluMetLeuThrGly 80  
 Db 185 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTACCTGACGAGAAATGTAAACGGGA 244  
 QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
 Db 245 GGTCTGCTTCCCTGGGCAACAGAGCCACCTGGGCGAGGAAAGTGCACG 295

## RESULT 2

US-09-738-973-442

Sequence 442, Application US/09738973

Patent No. US20020110563A1

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.

APPLICANT: Henderson, Robert A.

APPLICANT: Iodes, Michael J.

APPLICANT: Flinn, Steven P.

APPLICANT: Mohamath, Raedoh

APPLICANT: Algate, Paul A.

APPLICANT: Secrist, Heather

APPLICANT: Indrias, Carol Yoseph

APPLICANT: Benson, Darin R.

APPLICANT: Elliott, Mark

APPLICANT: Mannion, Jane

APPLICANT: Kalos, Michael D.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

FILE REFERENCE: 210121.475C9

CURRENT APPLICATION NUMBER: US/09/738,973

CURRENT FILING DATE: 2000-12-14

NUMBER OF SEQ ID NOS: 587

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 442

LENGTH: 337

TYPE: DNA

ORGANISM: Homo sapiens

US-09-738-973-442

## Alignment Scores:

Pred. No.: 3,54e-94 Length: 337  
 Score: 97.00 Matches: 97

Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-854-133-586 (1-97) x US-09-738-973-442 (1-337)

QY 1 GluValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 20  
 Db 5 GAGTTTAAAGTGAAGACAGATCATGCGCTGGGTGACAGTGAAGTCTGTCTCAACA 64  
 QY 21 GluLeuArgLysGlyGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 40  
 Db 65 GAATTAAGGAAAAAGAAAAAGAGAGAGAGAAATTCAGGCCAATGTGTGCG 124  
 QY 41 IleAspPheIlePheTrpIlePheTrpIleLeuLeuPheSerHisIleTrpIleGln 60  
 Db 125 ATGATTTTATCATATTCGTGATTTTGTGATTTCTTTTCTCATCATCGATTGCG 184  
 QY 61 GluSerLeuLeuGlyProPheSerProLysGluValThrCysArgGluMetLeuThrGly 80  
 Db 185 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTACCTGACGAGAAATGTAAACGGGA 244  
 QY 81 GlyCysLeuProTrpAlaThrArgSerHisLeuGlyArgArgLysCysSer 97  
 Db 245 GGTCTGCTTCCCTGGGCAACAGAGCCACCTGGGCGAGGAAAGTGCACG 295

## RESULT 3

US-09-854-133-441

Sequence 441, Application US/09854133

Publication No. US20020183499A1

GENERAL INFORMATION:

APPLICANT: Iodes, Michael J.

APPLICANT: Mohamath, Raedoh

APPLICANT: Henderson, Robert A.

APPLICANT: Benson, Darin R.

APPLICANT: Secrist, Heather

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

FILE REFERENCE: 210121.475C10

CURRENT APPLICATION NUMBER: US/09/854,133

CURRENT FILING DATE: 2001-05-11

NUMBER OF SEQ ID NOS: 735

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 441

LENGTH: 5981

TYPE: DNA

ORGANISM: Homo sapiens

US-09-854-133-441

## Alignment Scores:

Pred. No.: 5,6e-92 Length: 5981  
 Score: 96.00 Matches: 96  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.97% Indels: 0  
 Gaps: 0

US-09-854-133-586 (1-97) x US-09-854-133-441 (1-5981)

QY 2 ValGluValSerArgAspHisAlaSerLeuGlyAspSerGluThrLeuSerGlnThr 21  
 Db 3 GTTGAAGTGAAGACAGATCATGCGCTGGGTGACAGTGAAGTCTGTCTCAACA 62  
 QY 22 LeuArgLysGlyGluArgLysLysArgGluArgLysPheGlnAlaAsnCysGly 41  
 Db 63 TTAAGGAAAAAGAAAAAGAGAGAGAGAAATTCAGGCCAATGTGTGCG 122  
 QY 42 AspPheIleIlePheTrpIlePheTrpIleLeuLeuPheSerHisIleTrpIleGln 61  
 Db 123 GATTTTATCATATTCGTGATTTTGTGATTTCTTTTCTCATCATCGATTGCG 182  
 QY 62 SerLeuLeuGlyProPheSerProLysGluValThrCysArgGluMetLeuThrGly 81

```
Db 183 AGCGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGAAATCTTAACGGAGGC 242
QY 82 CysleuprotpralathArgSerHisleuGlyArgGlyCysSer 97
Db 243 TGCCTTCCCTGGGCAACAAGAGCCACTGGGACAGAGAAAGTGCAGC 290

RESULT 4
US-09-738-973-441
; Sequence 441, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Monamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 441
; LENGTH: 5981
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-738-973-441

Alignment Scores:
Pred. No.: 5,6e-92 Length: 5981
Score: 96.00 Matches: 96
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.97% Indels: 0
Gaps: 0
DB: 10

US-09-854-133-586 (1-97) x US-09-738-973-441 (1-5981)
QY 2 ValGluValSerArgSPHISAlaSerleuGlyAspSerGluThrleuSerGlnThrGlu 21
Db 3 GTTCAAGTGCAGCAGATCATGCCAGCGTGGTGACAGTGCAGCTCTGTCAACAGAA 62
QY 22 LeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGlyIle 41
Db 63 TTAAGGAAAAAAGAAAAAGAAAAAGAGAGAGAAATTCACGGCCAAATGTGGCATTA 122
QY 42 AspHeleIlePheTrpIlePheTrpIleleuPheSerHisStrpIleGlnIle 61
Db 123 GATTTATCATATCTGTGATTTTGTGATTTCTTTGTTTCTCATCTGATTCAGGAA 182
QY 62 SerleuLeuCysProPheSerProLysGluValThrCysArgGluMetleuThrGly 81
Db 183 AGCGTGTGTCACCATCTCCAAAGAGGTTACCTGCAGGAAATCTTAACGGAGGC 242
QY 82 CysleuprotpralathArgSerHisleuGlyArgGlyCysSer 97
Db 243 TGCCTTCCCTGGGCAACAAGAGCCACTGGGACAGAGAAAGTGCAGC 290

RESULT 5
US-09-854-133-440
; Sequence 440, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
```

```
; APPLICANT: Monamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 440
; LENGTH: 2239
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-854-133-440

Alignment Scores:
Pred. No.: 3,45e-89 Length: 2239
Score: 93.00 Matches: 93
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 95.88% Indels: 0
Gaps: 0
DB: 9

US-09-854-133-586 (1-97) x US-09-854-133-440 (1-2239)
QY 1 GluValGluValSerArgSPHISAlaSerleuGlyAspSerGluThrleuSerGlnThr 20
Db 2 GAGGTGAAGTGCAGCAGATCATGCCAGCGTGGTGACAGTGCAGTCTGTCTCAACA 61
QY 21 GluLeuArgLysLysGluArgLysLysLysArgGluArgLysPheGlnAlaAsnCysGly 40
Db 62 GAATTAGAAAAAAGAAAAAGAAAAAGAGAGAAATTCACGGCCAAATGTGGC 121
QY 41 IleAspHeleIlePheTrpIlePheTrpIleleuPheSerHisStrpIleGln 60
Db 122 ATAGATTTCATATCTGTGATTTTGTGATTTCTTTGTTTCTCATCTGATTCAG 181
QY 61 GluSerleuLeuCysProPheSerProLysGluValThrCysArgGluMetleuThrGly 80
Db 182 GAAGCGCTGTGTCTCCACCATCTCCAAAGAGGTTACCTGCAGGAAATGTAAACGGA 241
QY 81 GlyCysleuprotpralathArgSerHisleuGlyArg 93
Db 242 GGCTGCCCTTCCCTGGGCAACAAGAGCCACTGGGACAG 280

RESULT 6
US-09-738-973-440
; Sequence 440, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Monamath, Raodoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: THE THERAPY AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 440
; LENGTH: 2239
```

01 GLuSerLeuLeuCysProProSerProLysGluValThrCysAraGlnMetLeuThrGln 90

**FEATURE:**

NAME/KEY: misc\_feature

```
LOCATION: (1)...(572)
OTHER INFORMATION: n = A,T,C or G
US-09-878-178-1307

Alignment Scores:
Pred. No.: 2,166-30      Length: 572
Score: 37.00             Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.14%      Indels: 0
DB: 9                     Gaps: 0

US-09-854-133-586 (1-97) x US-09-878-178-1307 (1-572)
QY 61 GlusertleucysProProserProlysGluValThrCysArgGluMetLeuThrGly 80
DB 256 GAAGGCTGTGTGTCTCCACCATCTCCAAAGAGGTTACTGCGAGGAAATGTTAACGGGA 315
QY 81 GlyCysLeuProTTPalathrArgSerHisLeuGlyArgArgLysCysSer 97
DB 316 GGCTGCTTCTCCCTGGGCAACAGAGGACCTGGGCGAGGAAAGTGCAGC 366

RESULT 10
US-10-146-502-1307
Sequence 1307, Application US/10146502
Publication No. US20030069180A1
GENERAL INFORMATION:
APPLICANT: Jiang, Yugu
APPLICANT: Harlocker, Susan L.
APPLICANT: Secrist, Heather
APPLICANT: Wang, Aljun
APPLICANT: Stolk, John A.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
FILE REFERENCE: 210121.527C2
CURRENT APPLICATION NUMBER: US/10/146,502
CURRENT FILING DATE: 2002-05-14
NUMBER OF SEQ ID NOS: 2241
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1307
LENGTH: 572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc-feature
LOCATION: 9, 19, 461, 497, 500, 502
OTHER INFORMATION: n = A,T,C or G
US-10-146-502-1307

Alignment Scores:
Pred. No.: 2,166-30      Length: 572
Score: 37.00             Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.14%      Indels: 0
DB: 9                     Gaps: 0

US-09-854-133-586 (1-97) x US-10-146-502-1307 (1-572)
QY 61 GlusertleucysProProserProlysGluValThrCysArgGluMetLeuThrGly 80
DB 256 GAAGGCTGTGTGTCTCCACCATCTCCAAAGAGGTTACTGCGAGGAAATGTTAACGGGA 315
QY 81 GlyCysLeuProTTPalathrArgSerHisLeuGlyArgArgLysCysSer 97
DB 316 GGCTGCTTCTCCCTGGGCAACAGAGGACCTGGGCGAGGAAAGTGCAGC 366

RESULT 11
US-10-163-866-53
Sequence 53, Application US/10163866
Publication No. US20030027188A1
GENERAL INFORMATION:
APPLICANT: EXELIXIS, INC.

TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
FILE REFERENCE: EX02-080C
CURRENT APPLICATION NUMBER: US/10/163,866
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: US 60/296,076
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/328,605
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/338,733
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: US 60/357,253
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/357,600
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 1528
TYPE: DNA
ORGANISM: Homo sapiens
US-10-163-866-52

Alignment Scores:
Pred. No.: 5,286-30      Length: 1528
Score: 37.00             Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.14%      Indels: 0
DB: 9                     Gaps: 0

US-09-854-133-586 (1-97) x US-10-163-866-53 (1-1268)
QY 61 GlusertleucysProProserProlysGluValThrCysArgGluMetLeuThrGly 80
DB 30 GAAGGCTGTGTGTCTCCACCATCTCCAAAGAGGTTACTGCGAGGAAATGTTAACGGGA 89
QY 81 GlyCysLeuProTTPalathrArgSerHisLeuGlyArgArgLysCysSer 97
DB 90 GGCTGCTTCTCCCTGGGCAACAGAGGACCTGGGCGAGGAAAGTGCAGC 140

RESULT 12
US-10-163-866-52
Sequence 52, Application US/10163866
Publication No. US20030027188A1
GENERAL INFORMATION:
APPLICANT: EXELIXIS, INC.
TITLE OF INVENTION: SLIC7s AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE
FILE REFERENCE: EX02-080C
CURRENT APPLICATION NUMBER: US/10/163,866
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: US 60/296,076
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/328,605
PRIOR FILING DATE: 2001-10-10
PRIOR APPLICATION NUMBER: US 60/338,733
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: US 60/357,253
PRIOR FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/357,600
PRIOR FILING DATE: 2002-02-15
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 1528
TYPE: DNA
ORGANISM: Homo sapiens
US-10-163-866-52

Alignment Scores:
Pred. No.: 5,286-30      Length: 1528
Score: 37.00             Matches: 37
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 38.14%      Indels: 0
DB: 9                     Gaps: 0
```

Query Match: 38.14% Indels: 0  
DB: 9 Gaps: 0  
US-09-854-133-586 (1-97) x US-10-163-866-52 (1-1528)  
QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
DB 30 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACTGCGAGGGAATGTTAACGGGA 89  
QY 81 GlyCysLeuProTTPaIaThrArgSerHisLeuGlyArgGlyCysSer 97  
DB 90 GGCTGCTTCCCTGGGCAACAGAGGACCTGGGCGAGGAAAGTGCAGC 140  
RESULT 13  
US-10-163-866-33  
; Sequence 33, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 33  
; LENGTH: 1542  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-33  
Alignment Scores:  
Pred. No.: 5.33e-30 Length: 1542  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
Gaps: 0  
US-09-854-133-586 (1-97) x US-10-163-866-33 (1-1542)  
QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
DB 16 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACTGCGAGGGAATGTTAACGGGA 75  
QY 81 GlyCysLeuProTTPaIaThrArgSerHisLeuGlyArgGlyCysSer 97  
DB 76 GGCTGCTTCCCTGGGCAACAGAGGACCTGGGCGAGGAAAGTGCAGC 126  
RESULT 14  
US-10-163-866-30  
; Sequence 30, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10

; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; PRIOR FILING DATE: 2002-02-15  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 30  
; LENGTH: 1861  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-30  
Alignment Scores:  
Pred. No.: 6.33e-30 Length: 1861  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
Gaps: 0  
US-09-854-133-586 (1-97) x US-10-163-866-30 (1-1861)  
QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
DB 243 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACTGCGAGGGAATGTTAACGGGA 302  
QY 81 GlyCysLeuProTTPaIaThrArgSerHisLeuGlyArgGlyCysSer 97  
DB 303 GGCTGCTTCCCTGGGCAACAGAGGACCTGGGCGAGGAAAGTGCAGC 353  
RESULT 15  
US-10-163-866-34  
; Sequence 34, Application US/10163866  
; Publication No. US20030027188A1  
; GENERAL INFORMATION:  
; APPLICANT: EXELIXIS, INC.  
; TITLE OF INVENTION: SLCTs AS MODIFIERS OF THE P53 PATHWAY AND METHODS OF USE  
; FILE REFERENCE: EX02-080C  
; CURRENT APPLICATION NUMBER: US/10/163,866  
; PRIOR FILING DATE: 2002-06-05  
; PRIOR APPLICATION NUMBER: US 60/296,076  
; PRIOR FILING DATE: 2001-06-05  
; PRIOR APPLICATION NUMBER: US 60/328,605  
; PRIOR FILING DATE: 2001-10-10  
; PRIOR APPLICATION NUMBER: US 60/338,733  
; PRIOR FILING DATE: 2001-10-22  
; PRIOR APPLICATION NUMBER: US 60/357,253  
; PRIOR FILING DATE: 2002-02-15  
; PRIOR APPLICATION NUMBER: US 60/357,600  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 34  
; LENGTH: 2000  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-163-866-34  
Alignment Scores:  
Pred. No.: 6.76e-30 Length: 2000  
Score: 37.00 Matches: 37  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 38.14% Indels: 0  
Gaps: 0  
US-09-854-133-586 (1-97) x US-10-163-866-34 (1-2000)  
QY 61 GluSerLeuLeuCyProProSerProLysGluValThrCysArgGluMetLeuThrGly 80  
DB 143 GAAAGCCTGTGTGTCCACCATCTCCAAAGAGGTTACTGCGAGGGAATGTTAACGGGA 202



;; PRIOR FILING DATE: 1998-05-14  
;; NUMBER OF SEQ ID NOS: 811  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 216  
;; LENGTH: 294  
;; TYPE: DNA  
;; ORGANISM: Murine  
;; FEATURE:  
;; NAME/KEY: misc-feature  
;; LOCATION: (1)...(294)  
;; OTHER INFORMATION: n = A,T,C or G  
US-09-933-797-216

## Alignment Scores:

Pred. No.: 0.00416 Length: 294  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-933-797-216 (1-294)

Qy 23 Arglystysglnuarglystyslgsarglunrg 33  
Db 61 AGAAGAGAAAGAAAGAAAGAAAGAGAGAGAGA 93

## RESULT 20

US-10-091-504-2209  
;; Sequence 2209, Application US/10091504  
;; Publication No. US20030059908A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
;; FILE REFERENCE: PC007C1  
;; CURRENT APPLICATION NUMBER: US/10/091,504  
;; NUMBER OF SEQ ID NOS: 2442  
;; PRIOR FILING DATE: 2002-03-07  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 2209  
;; LENGTH: 32190  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-091-504-2209

## Alignment Scores:

Pred. No.: 0.302 Length: 32190  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-10-091-504-2209 (1-32190)

Qy 9 Alaserleuglyaspsergluthrleusergln 19  
Db 30055 GCCAGCCTGGCGACGACGACTGTCTCAA 30087

## RESULT 21

US-09-764-869-2209  
;; Sequence 2209, Application US/09764869  
;; Patent No. US20020061521A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Rosen et al.  
;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
;; FILE REFERENCE: PC007  
;; CURRENT APPLICATION NUMBER: US/09/764,869  
;; PRIOR FILING DATE: 2001-01-17  
;; NUMBER OF SEQ ID NOS: 2442  
;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 2209  
;; LENGTH: 32190  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-764-869-2209

## Alignment Scores:

Pred. No.: 0.302 Length: 32190  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-764-869-2209 (1-32190)

Qy 9 Alaserleuglyaspsergluthrleusergln 19  
Db 30055 GCCAGCCTGGCGACGACGACTGTCTCAA 30087

## RESULT 22

US-09-998-598-2201/c  
;; Sequence 2201, Application US/0998598  
;; Patent No. US20020150922A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Stolk, John A.  
;; APPLICANT: Xu, Jiangchun  
;; APPLICANT: Chenault, Ruth A.  
;; APPLICANT: Meagher, Madelein Joy  
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
;; FILE REFERENCE: 210121.361  
;; CURRENT APPLICATION NUMBER: US/09/998,598  
;; NUMBER OF SEQ ID NOS: 2606  
;; SOFTWARE: Corixa Invention Disclosure Database  
;; SEQ ID NO 2201  
;; LENGTH: 261  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-998-598-2201

## Alignment Scores:

Pred. No.: 0.0428 Length: 261  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
Gaps: 0

US-09-854-133-586 (1-97) x US-09-998-598-2201 (1-261)

Qy 10 Serleuglyaspsergluthrleusergln 19  
Db 104 ACCCTGGGTGACAGTGACCTGTCTCAA 75

## RESULT 23

US-09-803-719-71  
;; Sequence 71, Application US/09803719  
;; Publication No. US20030044783A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Williams, Lewis T.  
;; APPLICANT: Escobedo, Jaime  
;; APPLICANT: Innis, Michael A.  
;; APPLICANT: Garcia, Pablo Dominguez  
;; APPLICANT: Sudduth-Klinger, Julie  
;; APPLICANT: Reinhard, Christoph  
;; APPLICANT: Gliese, Klaus  
;; APPLICANT: Randazzo, Filippo  
;; APPLICANT: Kennedy, Giulio C.  
;; APPLICANT: Pot, David  
;; APPLICANT: Kassam, Altaf  
;; APPLICANT: Lamson, George



```
; SEQ ID NO 234
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-234

Alignment Scores:
Pred. No.: 0.0473 Length: 291
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
Gaps: 0
DB: 9

US-09-854-133-586 (1-97) x US-09-803-719-71 (1-291)
QY 10 SerLeuGlyAspSergIuThrIeuSergIn 19
DB 97 AGCCTGGGTGACGACGAGACTTATCTCAA 126

RESULT 24
US-09-803-719-234
; Sequence 234, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulio C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radomir
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 234
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-234

Alignment Scores:
Pred. No.: 0.0473 Length: 291
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
Gaps: 0
DB: 9

US-09-854-133-586 (1-97) x US-09-803-719-234 (1-291)
QY 10 SerLeuGlyAspSergIuThrIeuSergIn 19
DB 28 AGCCTGGGTGACGACGAGACTTATCTCAA 57

RESULT 25
US-09-803-719-83
; Sequence 83, Application US/09803719
; Publication No. US20030044783A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Escobedo, Jaime
; APPLICANT: Innis, Michael A.
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Sudduth-Klinger, Julie
; APPLICANT: Reinhard, Christoph
; APPLICANT: Giese, Klaus
; APPLICANT: Randazzo, Filippo
; APPLICANT: Kennedy, Giulio C.
; APPLICANT: Pot, David
; APPLICANT: Kassam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Drmanac, Radomir
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Human Genes and Gene Products
; FILE REFERENCE: 1624.002
; CURRENT APPLICATION NUMBER: US/09/803,719
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/188,609
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 2396
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 83
; LENGTH: 303
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-803-719-83

Alignment Scores:
Pred. No.: 0.049 Length: 303
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.31% Indels: 0
Gaps: 0
DB: 9

US-09-854-133-586 (1-97) x US-09-803-719-83 (1-303)
QY 10 SerLeuGlyAspSergIuThrIeuSergIn 19
DB 11 AGCCTGGGTGACGACGAGACTTATCTCAA 11
```

Db 97 AGCTGGGTGACAGCGAGACTTATCTCAA 126

RESULT 26

US-09-954-531-403/c

Sequence 403, Application US/09954531

Patent No. US20020165180A1

GENERAL INFORMATION:

APPLICANT: Weaver, Zoe

TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand

FILE REFERENCE: 689290-77

CURRENT APPLICATION NUMBER: US/09/954,531

CURRENT FILING DATE: 2002-05-02

PRIOR APPLICATION NUMBER: US/60/233,133

PRIOR FILING DATE: 2000-09-18

PRIOR APPLICATION NUMBER: US/60/234,009

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,034

PRIOR FILING DATE: 2000-09-20

PRIOR APPLICATION NUMBER: US/60/234,509

PRIOR FILING DATE: 2000-09-22

PRIOR APPLICATION NUMBER: US/60/234,567

NUMBER OF SEQ ID NOS: 1392

SOFTWARE: PatentIn version 3.0

SEQ ID NO 403

LENGTH: 351

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(351)

OTHER INFORMATION: n=a,t,g or c

US-09-954-531-403

Alignment Scores:

Pred. No.: 0.0561

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

Gaps: 0

Length: 351

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

US-09-854-133-586 (1-97) x US-09-954-531-403 (1-351)

QY 10 SerLeuGlyASpSerGIuThrLeuSerGIu 19

Db 153 AGCTGGGTGACAGCGAGACTTATCTCAA 124

RESULT 27

US-09-867-701-7120/c

Sequence 7120, Application US/09867701

Patent No. US20020132237A1

GENERAL INFORMATION:

APPLICANT: Aglate, Paul A.

APPLICANT: Jones, Robert

APPLICANT: Harlocker, Susan L.

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

FILE REFERENCE: 210121.497

CURRENT APPLICATION NUMBER: US/09/867,701

CURRENT FILING DATE: 2001-05-29

NUMBER OF SEQ ID NOS: 10912

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7120

LENGTH: 418

TYPE: DNA

ORGANISM: Homo sapien

US-09-867-701-7120

Alignment Scores:

Pred. No.: 0.0658

Score: 10.00

Length: 418

Matches: 10

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 10

Gaps: 0

US-09-854-133-586 (1-97) x US-09-867-701-7120 (1-418)

QY 10 SerLeuGlyASpSerGIuThrLeuSerGIu 19

Db 336 AGCTGGGTGACAGCGAGACTTATCTCAA 307

RESULT 28

US-09-918-995-12641

Sequence 12641, Application US/09918995

Publication No. US20030073623A1

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED

FILE REFERENCE: 20411-756

CURRENT APPLICATION NUMBER: US/09/918,995

CURRENT FILING DATE: 2001-07-30

PRIOR APPLICATION NUMBER: US/09/235,076

PRIOR FILING DATE: 1999-01-20

NUMBER OF SEQ ID NOS: 38054

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 12641

LENGTH: 483

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc\_feature

LOCATION: (1)...(483)

OTHER INFORMATION: n = A,T,C or G

US-09-918-995-12641

Alignment Scores:

Pred. No.: 0.0751

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

Gaps: 0

Length: 483

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

US-09-854-133-586 (1-97) x US-09-918-995-12641 (1-483)

QY 10 SerLeuGlyASpSerGIuThrLeuSerGIu 19

Db 423 AGCTGGGTGACAGCGAGACTTATCTCAA 452

RESULT 29

US-09-764-891-1995/c

Sequence 1995, Application US/09764891

Publication No. US20030077808A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC006

CURRENT APPLICATION NUMBER: US/09/764,891

CURRENT FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 10231

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 1995

LENGTH: 497

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (109)

OTHER INFORMATION: n equals a,t,g, or c

NAME/KEY: SITE

LOCATION: (489)



APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
TITLE OF INVENTION:  
FILE REFERENCE:  
CURRENT APPLICATION NUMBER: US/10/028,072  
CURRENT FILING DATE: 2001-12-19  
PRIOR APPLICATION NUMBER: 60/049911  
PRIOR FILING DATE: 1997-06-18  
PRIOR APPLICATION NUMBER: 60/056974  
PRIOR FILING DATE: 1997-08-26  
PRIOR APPLICATION NUMBER: 60/059113  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059115  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059117  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059122  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059184  
PRIOR FILING DATE: 1997-09-17  
PRIOR APPLICATION NUMBER: 60/059263  
PRIOR FILING DATE: 1997-09-18  
PRIOR APPLICATION NUMBER: 60/059352  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059588  
PRIOR FILING DATE: 1997-09-19  
PRIOR APPLICATION NUMBER: 60/059836  
PRIOR FILING DATE: 1997-09-24  
PRIOR APPLICATION NUMBER: 60/062250  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062285  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062287  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/062814  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/062816  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063045  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063082  
PRIOR FILING DATE: 1997-10-31  
PRIOR APPLICATION NUMBER: 60/063127  
PRIOR FILING DATE: 1997-10-24  
PRIOR APPLICATION NUMBER: 60/063327  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063329  
PRIOR FILING DATE: 1997-10-27  
PRIOR APPLICATION NUMBER: 60/063550  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063561  
PRIOR FILING DATE: 1997-10-28  
PRIOR APPLICATION NUMBER: 60/063704  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063733  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063735  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063758  
PRIOR FILING DATE: 1997-10-29  
PRIOR APPLICATION NUMBER: 60/063755  
PRIOR FILING DATE: 1997-10-17  
PRIOR APPLICATION NUMBER: 60/064248  
PRIOR FILING DATE: 1997-11-03  
PRIOR APPLICATION NUMBER: 60/064809  
PRIOR FILING DATE: 1997-11-07

PRIOR APPLICATION NUMBER: 60/065186  
PRIOR FILING DATE: 1997-11-12  
PRIOR APPLICATION NUMBER: 60/065846  
PRIOR FILING DATE: 1997-11-17  
PRIOR APPLICATION NUMBER: 60/066364  
PRIOR FILING DATE: 1997-11-21  
PRIOR APPLICATION NUMBER: 60/066453  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066511  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/066770  
PRIOR FILING DATE: 1997-11-24  
PRIOR APPLICATION NUMBER: 60/069212  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069278  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069334  
PRIOR FILING DATE: 1997-12-11  
PRIOR APPLICATION NUMBER: 60/069694  
PRIOR FILING DATE: 1997-12-16  
PRIOR APPLICATION NUMBER: 60/072320  
PRIOR FILING DATE: 1998-01-23  
PRIOR APPLICATION NUMBER: 60/073612  
PRIOR FILING DATE: 1998-02-04  
PRIOR APPLICATION NUMBER: 60/074086  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/074092  
PRIOR FILING DATE: 1998-02-09  
PRIOR APPLICATION NUMBER: 60/077791  
PRIOR FILING DATE: 1998-03-12  
PRIOR APPLICATION NUMBER: 60/078910  
PRIOR FILING DATE: 1998-03-20  
PRIOR APPLICATION NUMBER: 60/079294  
PRIOR FILING DATE: 1998-03-25  
PRIOR APPLICATION NUMBER: 60/079663  
PRIOR FILING DATE: 1998-02-27  
PRIOR APPLICATION NUMBER: 60/079728  
PRIOR FILING DATE: 1998-03-27  
PRIOR APPLICATION NUMBER: 60/080165  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 60/081203  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081229  
PRIOR FILING DATE: 1998-04-09  
PRIOR APPLICATION NUMBER: 60/081695  
PRIOR FILING DATE: 1998-04-14  
PRIOR APPLICATION NUMBER: 60/081817  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/081818  
PRIOR FILING DATE: 1998-04-15  
PRIOR APPLICATION NUMBER: 60/082999  
PRIOR FILING DATE: 1998-04-24  
PRIOR APPLICATION NUMBER: 60/083322  
PRIOR FILING DATE: 1998-04-28  
PRIOR APPLICATION NUMBER: 60/083545  
PRIOR FILING DATE: 1998-04-29  
PRIOR APPLICATION NUMBER: 60/084600  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084627  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/084637  
PRIOR FILING DATE: 1998-05-07  
PRIOR APPLICATION NUMBER: 60/085149  
PRIOR FILING DATE: 1998-05-12  
PRIOR APPLICATION NUMBER: 60/085323  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085338  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085339  
PRIOR FILING DATE: 1998-05-13  
PRIOR APPLICATION NUMBER: 60/085579  
PRIOR FILING DATE: 1998-05-15  
PRIOR APPLICATION NUMBER: 60/085697

```
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086414
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086430
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087106
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/088026
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088730
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088741
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088810
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088858
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089599
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089907
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/089947
;; PRIOR FILING DATE: 1998-06-19
;; PRIOR APPLICATION NUMBER: 60/090349
;; PRIOR FILING DATE: 1998-06-23
;; PRIOR APPLICATION NUMBER: 60/090429
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090445
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090538
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090863
;; PRIOR FILING DATE: 1998-06-26
;; PRIOR APPLICATION NUMBER: 60/091360
;; PRIOR FILING DATE: 1998-07-01
;; PRIOR APPLICATION NUMBER: 60/091519
;; PRIOR FILING DATE: 1998-07-02
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07

Alignment Scores:
Pred. No.: 0.202      Length: 1432
Score: 10.00         Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 10.31%      Indels: 0
DB: 9                Gaps: 0

US-09-854-133-586 (1-97) x US-10-028-072-255 (1-1432)
QY 10 Serleuglyaspsergluthreusergln 19
Db 1261 AGCCTGGTGACGACGAGACCCGTCTCA 1290

RESULT 33
US-10-121-049-255
; Sequence 255, Application US/10121049
; Publication No. US20030022239A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
```

```
;; APPLICANT: Sherwood, Steven
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stewart, Timothy A.
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Watanabe, Colin K
;; APPLICANT: Wood, William
;; APPLICANT: Zhang, Zemin
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P33081C17
;; CURRENT APPLICATION NUMBER: US/10/121,049
;; CURRENT FILING DATE: 2002-04-12
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 550
;; SEQ ID NO 255
;; LENGTH: 1432
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-121-049-255

Alignment Scores:
Pred. No.: 0.202      Length: 1432
Score: 10.00         Matches: 10
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 10.31%      Indels: 0
DB: 9                Gaps: 0

US-09-854-133-586 (1-97) x US-10-121-049-255 (1-1432)
QY 10 Serleuglyaspsergluthreusergln 19
Db 1261 AGCCTGGTGACGACGAGACCCGTCTCA 1290

RESULT 34
US-10-123-904-255
; Sequence 255, Application US/10133904
; Publication No. US20030022328A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C54
; CURRENT APPLICATION NUMBER: US/10/123,904
; CURRENT FILING DATE: 2002-04-16
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 255
; LENGTH: 1432
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-123-904-255

Alignment Scores:
Pred. No.: 0.202      Length: 1432
Score: 10.00         Matches: 10
Percent Similarity: 100.00%      Conservative: 0
```

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-123-904-255 (1-1432)

QY 10 SerLeuGIyAspSerGIuThIleuSerGI 19

DB 1261 AGCCTGGGTGACAGCGAGACCCCTGTCTCA 1290

RESULT 35

US-10-140-470-255

Sequence 255, Application US/10140470

Publication No. US2003002231A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3330R1C160

CURRENT APPLICATION NUMBER: US/10/140,470

Prior Application removed - See File Wrapper

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 255

LENGTH: 1432

TYPE: DNA

ORGANISM: Homo Sapien

US-10-140-470-255

Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

Gaps: 0

Length: 1432

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-854-133-586 (1-97) x US-10-140-470-255 (1-1432)

QY 10 SerLeuGIyAspSerGIuThIleuSerGI 19

DB 1261 AGCCTGGGTGACAGCGAGACCCCTGTCTCA 1290

RESULT 36

US-10-175-746-255

Sequence 255, Application US/10175746

Publication No. US20030027270A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3330R1C382

CURRENT APPLICATION NUMBER: US/10/176,918

Prior Application removed - See File Wrapper or Palm

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3330R1C353

CURRENT APPLICATION NUMBER: US/10/175,746

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 255

LENGTH: 1432

TYPE: DNA

ORGANISM: Homo Sapien

US-10-175-746-255

Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 10.31%

DB: 9

Gaps: 0

Length: 1432

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

US-09-854-133-586 (1-97) x US-10-175-746-255 (1-1432)

QY 10 SerLeuGIyAspSerGIuThIleuSerGI 19

DB 1261 AGCCTGGGTGACAGCGAGACCCCTGTCTCA 1290

RESULT 37

US-10-176-918-255

Sequence 255, Application US/10176918

Publication No. US20030027275A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Gao, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P3330R1C382

CURRENT APPLICATION NUMBER: US/10/176,918

Prior Application removed - See File Wrapper or Palm

NUMBER OF SEQ ID NOS: 550

SEQ ID NO 255

LENGTH: 1432

TYPE: DNA

ORGANISM: Homo Sapien

US-10-176-918-255

Alignment Scores:

Pred. No.: 0.202

Score: 10.00

Length: 1432

Matches: 10

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-176-921-255 (1-1432)

OY 10 SerLeuGIYAspSerGIuThrLeuSerGI 19  
1261 AGCCTGGGTGACAGGACGACCTGTCTCAA 1290

RESULT 38

US-10-176-921-255  
Sequence 255, Application US/10176921  
Publication No. US20030027276A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C288  
CURRENT APPLICATION NUMBER: US/10/176,921  
CURRENT FILING DATE: 2002-06-20  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 255  
LENGTH: 1432  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-176-921-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-176-921-255 (1-1432)

OY 10 SerLeuGIYAspSerGIuThrLeuSerGI 19  
1261 AGCCTGGGTGACAGGACGACCTGTCTCAA 1290

RESULT 39

US-10-137-865-255  
Sequence 255, Application US/10137865  
Publication No. US20030032155A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C154  
CURRENT APPLICATION NUMBER: US/10/137,865  
CURRENT FILING DATE: 2002-05-03  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 255  
LENGTH: 1432  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-137-865-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-137-865-255 (1-1432)

OY 10 SerLeuGIYAspSerGIuThrLeuSerGI 19  
1261 AGCCTGGGTGACAGGACGACCTGTCTCAA 1290

RESULT 40

US-10-140-474-255  
Sequence 255, Application US/10140474  
Publication No. US20030032156A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: Deforge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P3330R1C162  
CURRENT APPLICATION NUMBER: US/10/140,474  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See Palm or File Wrapper  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 255  
LENGTH: 1432  
TYPE: DNA  
ORGANISM: Homo Sapien  
US-10-140-474-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432

Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-140-474-255 (1-1432)

OY 10 SerLeuGIyaspSergIurhIeusSergIn 19

DB 1261 AGCCTGGTGACAGCAGACCCGTCTCA 1290

RESULT 41

US-10-142-431-255  
; Sequence 255, Application US/10142431  
; Publication No. US20030036179A1  
; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Beresini, Maureen  
;; APPLICANT: DeForge, Laura  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Paul J.  
;; APPLICANT: Godowski, Audrey  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Sherwood, Steven  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Watanabe, Colin K  
;; APPLICANT: Wood, William  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3330R1C251  
;; CURRENT APPLICATION NUMBER: US/10/142,431  
;; PRIOR APPLICATION REMOVED - See File Wrapper or Palm  
;; NUMBER OF SEQ ID NOS: 550  
;; SEQ ID NO 255  
;; LENGTH: 1432  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-142-431-255

Alignment Scores:

Pred. No.: 0.202 Length: 1432  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 10.31% Indels: 0  
DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-142-431-255 (1-1432)

OY 10 SerLeuGIyaspSergIurhIeusSergIn 19

DB 1261 AGCCTGGTGACAGCAGACCCGTCTCA 1290

RESULT 42

US-10-143-114-255

; Sequence 255, Application US/10143114  
; Publication No. US20030036180A1  
; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Beresini, Maureen  
;; APPLICANT: DeForge, Laura  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerritsen, Mary E.

;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Sherwood, Steven  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Watanabe, Colin K  
;; APPLICANT: Wood, William  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3330R1C211  
;; CURRENT APPLICATION NUMBER: US/10/143,114  
;; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
;; NUMBER OF SEQ ID NOS: 550  
;; SEQ ID NO 255  
;; LENGTH: 1432  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-143-114-255

US-09-854-133-586 (1-97) x US-10-143-114-255 (1-1432)

OY 10 SerLeuGIyaspSergIurhIeusSergIn 19

DB 1261 AGCCTGGTGACAGCAGACCCGTCTCA 1290

RESULT 43

US-10-140-002-255

; Sequence 255, Application US/10140002  
; Publication No. US20030037623A1  
; GENERAL INFORMATION:

;; APPLICANT: Baker, Kevin P.  
;; APPLICANT: Beresini, Maureen  
;; APPLICANT: DeForge, Laura  
;; APPLICANT: Desnoyers, Luc  
;; APPLICANT: Filvaroff, Ellen  
;; APPLICANT: Gao, Wei-Qiang  
;; APPLICANT: Gerritsen, Mary E.  
;; APPLICANT: Goddard, Audrey  
;; APPLICANT: Godowski, Paul J.  
;; APPLICANT: Gurney, Austin L.  
;; APPLICANT: Sherwood, Steven  
;; APPLICANT: Smith, Victoria  
;; APPLICANT: Stewart, Timothy A.  
;; APPLICANT: Tumas, Daniel  
;; APPLICANT: Watanabe, Colin K  
;; APPLICANT: Wood, William  
;; APPLICANT: Zhang, Zemin  
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
;; FILE REFERENCE: P3330R1C39  
;; CURRENT APPLICATION NUMBER: US/10/140,002  
;; PRIOR APPLICATION REMOVED - See Palm or File Wrapper  
;; NUMBER OF SEQ ID NOS: 550  
;; SEQ ID NO 255  
;; LENGTH: 1432  
;; TYPE: DNA  
;; ORGANISM: Homo Sapien  
US-10-140-002-255

Alignment Scores:



Pred. No.: 0.202 Length: 1432  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-140-002-255 (1-1432)

OY 10 SerLeuGLyAspSerGluThrLeuSerGln 19

DB 1261 AGCCTGGGTGACAGCGAGACCTGTCTCAA 1290

RESULT 44

US-10-142-419-255

Sequence 255, Application US/10142419

Publication No. US20030044945A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura

APPLICANT: Desnoyers, Luc

APPLICANT: Filvaroff, Ellen

APPLICANT: Geo, Wei-Qiang

APPLICANT: Gerritsen, Mary E.

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Sherwood, Steven

APPLICANT: Smith, Victoria

APPLICANT: Stewart, Timothy A.

APPLICANT: Tumas, Daniel

APPLICANT: Watanabe, Colin K

APPLICANT: Wood, William

APPLICANT: Zhang, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Zhen, Zemin

APPLICANT: Gerritsen, Mary E.  
 APPLICANT: Goddard, Audrey  
 APPLICANT: Godowski, Paul J.  
 APPLICANT: Gurney, Austin L.  
 APPLICANT: Sherwood, Steven  
 APPLICANT: Smith, Victoria  
 APPLICANT: Stewart, Timothy A.  
 APPLICANT: Tumas, Daniel  
 APPLICANT: Watanabe, Colin K  
 APPLICANT: Wood, William  
 APPLICANT: Zhang, Zemin  
 APPLICANT: Zhen, Zemin  
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 FILE REFERENCE: P3330R1C38  
 CURRENT APPLICATION NUMBER: US/10/123,262  
 CURRENT FILING DATE: 2002-04-15  
 Prior Application removed - see file wrapper or palm  
 NUMBER OF SEQ ID NOS: 550  
 SEQ ID NO 255  
 LENGTH: 1432  
 TYPE: DNA  
 ORGANISM: Homo Sapien  
 US-10-123-262-255

Alignment Scores:  
 Pred. No.: 0.202 Length: 1432  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 9 Gaps: 0

US-09-854-133-586 (1-97) x US-10-123-262-255 (1-1432)

OY 10 SerLeuGLyAspSerGluThrLeuSerGln 19

DB 1261 AGCCTGGGTGACAGCGAGACCTGTCTCAA 1290

Search completed: May 11, 2003, 18:45:54  
 Job time: 133.027 secs



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## SUMMARIES

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: May 11, 2003, 16:29:05 ; Search time 1473.03 Seconds  
(without alignments)

1066.486 Million cell updates/sec

Title: US-09-854-133-586

Perfect score: 97  
Sequence: 1 EVEVSRDHASLGDESLISQT.....LTGCLPWATNSHLGRKCS 97

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Word size: 1

Total number of hits satisfying chosen parameters: 32308013

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 60 summaries

Command line parameters:  
-MODE=frame+ p2n.model -DEV=xlh  
-Q/cgn2.1/USPTO.spool/US09854133/runat.05052003.174133.718/app.query.fasta.1.462  
-DB=EST -OEMT=fastap -SUFFIX=oligo.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=60  
-DOCALIGN=200 -PWR\_SCORE=quality -THR\_MIN=1 -ALIGN=45 -MODE=LOCAL -OUTPMT=ptc  
-NORM=ext -HEAPSIZ=500 -MNTEN=0 -MAXLEN=2000000000  
-USER=US09854133@cgn2.1.2013@runat.05052003.174133.718 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT=120  
-MARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6 -Fgapext=7  
-YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

EST.\*  
1: em\_estba:\*  
2: em\_estnum:\*  
3: em\_estlin:\*  
4: em\_estinu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Result No.	Score	Match	Length	DB ID	Description
1	37	38.1	936	12	BG284503
2	37	38.1	1072	12	BG388107
3	33	34.0	910	12	BG326527
4	21	21.6	557	17	AQ002318
5	19	19.6	123	9	A1313891
6	15	15.5	350	12	BF095483
7	13	13.4	1184	12	BF526812
8	12	12.4	357	17	AZ020264
9	12	12.4	589	17	AZ799582
10	12	12.4	629	17	AQ351658
11	12	12.4	629	17	AQ352329
12	12	12.4	664	17	AZ258055
13	12	12.4	871	17	AZ747681
14	12	12.4	882	12	BE880444
15	12	12.4	1032	12	BM557601
16	11	11.3	210	12	BF809298
17	11	11.3	227	12	BF743959
18	11	11.3	227	12	BF818399
19	11	11.3	259	17	AZ278638
20	11	11.3	374	14	B0558860
21	11	11.3	382	9	A1628395
22	11	11.3	442	13	BM196350
23	11	11.3	445	17	BM117011
24	11	11.3	454	17	AQ014518
25	11	11.3	464	17	AZ356753
26	11	11.3	495	17	AZ779079
27	11	11.3	500	17	BH858364
28	11	11.3	506	17	AZ750540
29	11	11.3	530	14	BQ086880
30	11	11.3	583	17	AQ347131
31	11	11.3	587	14	BQ086649
32	11	11.3	643	17	AG075749
33	11	11.3	652	17	AZ791667
34	11	11.3	677	17	AG126279
35	11	11.3	689	17	AQ947014
36	11	11.3	696	13	BI912987
37	11	11.3	917	12	BE733566
38	11	11.3	1024	12	BE735002
39	11	11.3	1036	13	BI912939
40	11	11.3	1061	17	BC010121
41	10	10.3	1113	13	AZ742823
42	10	10.3	1113	12	BF853221
43	10	10.3	1127	12	BF853221
44	10	10.3	1127	16	BI122746
45	10	10.3	1155	10	BB325488
46	10	10.3	157	10	BB323023
47	10	10.3	158	10	BB340999
48	10	10.3	159	10	BB340999
49	10	10.3	160	10	BB323163
50	10	10.3	162	10	BB182002
51	10	10.3	162	10	BB328913
52	10	10.3	163	10	BB328913
53	10	10.3	163	14	BM940420
54	10	10.3	164	10	BB323785
55	10	10.3	164	10	BB323785
56	10	10.3	165	10	BB323785
57	10	10.3	179	10	BB324227
58	10	10.3	181	10	BB323547
59	10	10.3	181	10	BB323799
60	10	10.3	182	9	AU069480

## ALIGNMENTS

US-09-854-133-586  
EVEVSRDHASLGDESLISQT.....LTGCLPWATNSHLGRKCS  
BG284503 602408645  
BG388107 602413070  
BG326527 602425373  
AQ002318 CTT-HSP-2  
A1313891 hpi-7 PMA  
BF095483 IL2-UT007  
BF526812 602070364  
AZ020264 RPCI-23-2  
AZ799582 2M0057F09  
AQ351658 CITR1-E1  
AQ352329 CITR1-E1  
AZ258055 RPCI-23-1  
AZ747681 RPCI-24-6  
BE880444 601491430  
BM557601 AGENCOURT  
BF809298 OYO-C1019  
BF743959 CM2-BT082  
BF818399 601816096  
AZ278638 RPCI-23-1  
B0558860 H4054603  
A1628395 ty76a03.x  
BM196350 C0327E04-  
BM117011 RPCI-24-3  
AQ014518 CTT-HSP-2  
AZ356753 LM0097P22  
AZ779079 2M0014P19  
BH858364 B5-164a2-  
AZ750540 RPCI-24-1  
BQ086880 LH140505.Y  
AQ347131 RPCI-11-4  
BQ086649 pan rtogl  
AG075749 pan rtogl  
AZ791667 2M0041K09  
AG126279 pan rtogl  
AQ947014 Sheared D  
BI912987 603176482  
BE733566 601565360  
BE735002 601565360  
BI912939 603176781  
BC010121 Homo sapi  
AZ742823 RPCI-24-1  
BI060988 IL3-UT011  
BF853221 MR2-END09  
AQ239603 CTT-HSP-2  
BI122746  
BB325488 BB325488  
BB323023 BB323023  
BB340999 BB340999  
BB323163 BB323163  
BB182002 BB182002  
BB328913 BB328913  
BB329139 BB329139  
BM940420 UT-M-CG0P  
BB323785 BB323785  
BB323479 BB323479  
BB323232 BB323232  
BB324227 BB324227  
BB323547 BB323547  
BB323799 BB323799  
AU069480 AU069480

```

RESULT 1
BG284503          936 bp      mRNA      linear      EST 21-FEB-2001
DEFINITION       602408645F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4537810 5',
ACCESSION        BG284503
VERSION          BG284503.1 GI:13035516
KEYWORDS
SOURCE           human.
ORGANISM         Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE        1 (bases 1 to 936)
AUTHORS          NIH-MGC http://mgc.ncl.nih.gov/.
TITLE            NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL          National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Life Technologies, Inc.
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L1AM10463 row: 1 column: 11
                  High quality sequence stop: 795.
FEATURES
SOURCE           Location/Qualifiers
                  1..936
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:4537810"
                  /clone_1db="NIH_MGC_91"
                  /tissue_type="adenocarcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: prostate; Vector: PCMV-SPORT6; Site_1: NotI;
                  Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                  Average insert size 1.4 kb. Library enriched for
                  full-length clones and constructed by Life Technologies.
                  Note: this is a NIH_MGC Library."
BASE COUNT       248 a      202 c      251 g      235 t
ORIGIN
Alignment Scores:
Pred. No.:      2,35e-26      Length:      936
Score:          37.00         Matches:      37
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    38.14%       Indels:        0
DB:             12           Gaps:          0
US-09-854-133-586 (1-97) x BG284503 (1-936)
QY 61 GluSerLeuLeuGlyProPserProLySGluValThrCysArgGluMetLeuThrGly 80
Db 278 GAAAGCCTGTGTGTCACCATCTCCAAAGAGAGGTTACCTGCGAAGAAATGTAAACGGGA 337
QY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyAlaArgIysCysSer 97
Db 338 GGCTGCTTCCTCGGCAACAGAGACCACTGGCAGAGAAAGTGCACG 388
RESULT 2
BG388107          1072 bp      mRNA      linear      EST 12-MAR-2001
DEFINITION       602413070F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:452136 5',
ACCESSION        BG388107
VERSION          BG388107.1 GI:13281553
KEYWORDS
SOURCE           human.
ORGANISM         Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

REFERENCE        1 (bases 1 to 1072)
AUTHORS          NIH-MGC http://mgc.ncl.nih.gov/.
TITLE            NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL          National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: ATCC
                  CDNA Library Preparation: Life Technologies, Inc.
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:
                  http://image.llnl.gov
                  Plate: L1AM10421 row: k column: 17
                  High quality sequence stop: 625.
FEATURES
SOURCE           Location/Qualifiers
                  1..1072
                  /organism="Homo sapiens"
                  /db_xref="taxon:9606"
                  /clone="IMAGE:452136"
                  /clone_1db="NIH_MGC_92"
                  /tissue_type="embryonal carcinoma, cell line"
                  /lab_host="DH10B (phage-resistant)"
                  /note="Organ: testis; Vector: PCMV-SPORT6; Site_1: NotI;
                  Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                  Average insert size 2.5 kb. Library enriched for
                  full-length clones and constructed by Life Technologies.
                  Note: this is a NIH_MGC Library."
BASE COUNT       331 a      229 c      270 g      242 t
ORIGIN
Alignment Scores:
Pred. No.:      2,61e-26      Length:      1072
Score:          37.00         Matches:      37
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:      0
Query Match:    38.14%       Indels:        0
DB:             12           Gaps:          0
US-09-854-133-586 (1-97) x BG388107 (1-1072)
QY 61 GluSerLeuLeuGlyProPserProLySGluValThrCysArgGluMetLeuThrGly 80
Db 296 GAAAGCCTGTGTGTCACCATCTCCAAAGAGAGGTTACCTGCGAAGAAATGTAAACGGGA 355
QY 81 GlyCysLeuProTrrPalatThrArgSerHisLeuGlyAlaArgIysCysSer 97
Db 356 GGCTGCTTCCTCGGCAACAGAGACCACTGGCAGAGAAAGTGCACG 406
RESULT 3
BG326527          910 bp      mRNA      linear      EST 27-FEB-2001
DEFINITION       602425373F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4562994 5',
ACCESSION        BG326527
VERSION          BG326527.1 GI:13132964
KEYWORDS
SOURCE           human.
ORGANISM         Homo sapiens
                  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE        1 (bases 1 to 910)
AUTHORS          NIH-MGC http://mgc.ncl.nih.gov/.
TITLE            NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL          National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT          Contact: Robert Strausberg, Ph.D.
                  Email: cgapbs-remail.nih.gov
                  Tissue Procurement: DCTD/DTF
                  CDNA Library Preparation: Ling Hong/Rubin Laboratory
                  DNA Sequencing by: Incyte Genomics, Inc.
                  Clone distribution: MGC clone distribution information can be
                  found through the I.M.A.G.E. Consortium/LLNL at:

```

http://image.jnl.gov  
Plate: LCM1276 row: b column: 19  
High quality sequence stop: 706.  
Location/Qualifiers

## FEATURES

## SOURCE

1..910  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4562994"  
/clone\_1lb="NIH\_MGC\_14"  
/issue\_type="renal cell adenocarcinoma"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: Kidney; Vector: pOT7; Site\_1: XhoI; Site\_2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 237 a 178 c 249 g 246 t  
ORIGIN

## Alignment Scores:

Pred. No.:	2..01e-22	Length:	910
Score:	33.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	34.02%	Indels:	0
DB:	12	Gaps:	0

US-09-854-133-586 (1-97) x BG326527 (1-910)

QY 65 CysProPterProlysgluValThrCysArgGluMetLeuThrGlyGlyCysLeuPro 84  
|||||  
Db 179 TGCCACCAATCTCCAAAGGAGGTACCTGACGGAATGTTAACGGAGGCTGCTTCC 238  
|||||  
QY 85 TrpAlaThrArgSerHisLeuGlyArgArgLyGlyCysSer 97  
|||||  
Db 239 TGGGCACACAGAGGCCACCTGGGACAGAGAAATGTCAGC 277  
|||||

RESULT 4  
A0002318 557 bp DNA linear GSS 26-JUN-1998  
LOCUS A0002318  
DEFINITION CIT-HSP-2283E8.TF CIT-HSP Homo sapiens genomic clone 2283E8, DNA  
SEQUENCE A0002318  
VERSION A0002318.1 GI:3029522  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 557)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,  
Simon,M. and Venter,J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
Unpublished (1998)  
Other GSSs: CIT-HSP-2283E8.TF  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: madams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21;  
Class: BAC ends.  
Location/Qualifiers

## FEATURES

## SOURCE

1..557  
/organism="Homo sapiens"  
/db\_xref="GDB:7147907"  
/db\_xref="taxon:9606"  
/clone="2283E8"  
/clone\_1lb="CIT-HSP"  
/sex="Male"  
/cell\_type="Sperm"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2: HindIII"  
HindIII

BASE COUNT 142 a 152 c 133 g 130 t  
ORIGIN

## Alignment Scores:

Pred. No.:	9.18e-11	Length:	557
Score:	21.00	Matches:	21
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	21.65%	Indels:	0
DB:	17	Gaps:	0

US-09-854-133-586 (1-97) x A0002318 (1-557)

QY 72 ValThrCysArgGluMetLeuThrGlyGlyCysLeuProTrpAlaThrArgSerHisLeu 91  
|||||  
Db 344 GTTACCTGCGAGGGAATGTTACGGAGGCTGCTTCCCTGGGCACACAGACCACTG 265  
|||||  
QY 92 G1Y 92  
|||  
Db 284 GGC 282

RESULT 5  
A1313891 123 bp mRNA linear EST 17-DEC-1998  
LOCUS A1313891  
DEFINITION hpi-7 PMA-induced HL60 cell subtraction library Homo sapiens CDNA,  
mRNA sequence.  
A1313891  
VERSION A1313891.1 GI:4029010  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 123)  
Hiken,J.F. and Huang,H.V.  
Genes Up-Regulated During PMA-Induced HL60 Cell Macrophage-Like  
Differentiation  
Unpublished (1998)  
Contact: Hiken JF  
Department of Molecular Microbiology  
Washington University School of Medicine  
660 South Euclid Ave, Saint Louis, MO 63110-1093, USA  
Tel: 314 362 2756  
Fax: 314 362 1232  
Email: hiken@borcim.wustl.edu  
Orientation of insert unknown.  
Location/Qualifiers

1..123  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_1lb="PMA-Induced HL60 cell subtraction library"  
/cell\_type="macrophage-like"  
/cell\_line="PMA-Induced HL60 human leukemic cell line"  
/lab\_host="MC1061"  
/note="Vector: pSPORT 1; Site\_1: EcoRI; Site\_2: EcoRI; RNA  
from uninduced HL60 cells was subtracted against RNA  
from HL60 cells that were induced 8h with 16 nM PMA using the  
method of Wang and Brown (Proc Natl Acad Sci USA 86: 11505  
)."  
Location/Qualifiers

## FEATURES

## SOURCE

BASE COUNT 23 a 37 c 34 g 29 t  
ORIGIN

## Alignment Scores:

Pred. No.: 2,65e-09  
 Score: 19.00  
 Percent Similarity: 100.00%  
 Best Local Similarity: 100.00%  
 Query Match: 19.59%  
 DB: 9  
 Matches: 123  
 Conserved: 19  
 Mismatches: 0  
 Indels: 0  
 Gaps: 0

US-09-854-133-586 (1-97) x A1313891 (1-123)

QY 78 LeuthrGlyGlyCysLeuProTrrPalatPharGserHisLeuGlyAraGlyCys 96  
 Db 59 TTAACGGAGGCTGCTCCCTCCGACACAGAGCCACCTGCGAGGAGAAAGTCC 3

RESULT 6  
 BF095483

LOCUS 350 bp mRNA linear EST 19-OCT-2000  
 DEFINITION IT2-UT0074-040900-152-C01 UT0074 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BF095483  
 VERSION BF095483.1 GI:10901193  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 350)  
 Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
 Nagai, M.A., da Silva, W. Jr., Zaglo, M.A., Bordin, S., Costa, F.F.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,  
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
 Simpson, A.J.G.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE

JOURNAL  
 MEDLINE  
 COMMENT  
 Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=6t2-IL2-UT0074-040  
 Seq primer: puc 18 forward  
 High quality sequence start: 24  
 High quality sequence stop: 350.  
 Location/Qualifiers

FEATURES  
 source  
 1..350  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="UT0074"  
 /dev\_stage="Adult"  
 /note="Organ: uterus; tumor; Vector: puc18; Site:1: Smal;  
 Site:2: SmaI; A mini-library was made by cloning products  
 derived from ORESTES PCR (U.S. Letters Patent application  
 No. 196,716 - Ludwig Institute for Cancer Research)  
 profiles into the puc 18 vector. Reverse transcription of  
 tissue mRNA and cDNA amplification were performed under  
 low stringency conditions."

BASE COUNT 95 a 80 c 87 t

Alignment Scores:

Pred. No.: 5.23e-05 Length: 350  
 Score: 15.00 Matches: 15  
 Percent Similarity: 100.00% Conserved: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 15.46% Indels: 0  
 DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BF095483 (1-350)

QY 61 GluserLeuGlyCysProPserProGlyGlyAlaThrCysArg 75  
 Db 297 GAAAGCCGTGTGTGTCCACATCTCCAAAGAGGTTACTCGACG 341

RESULT 7  
 BF526812

LOCUS 1184 bp mRNA linear EST 11-DEC-2000  
 DEFINITION 602070364F1 NCI CGAP\_Brn64 Homo sapiens cDNA clone IMAGE:4213479  
 5', mRNA sequence.  
 ACCESSION BF526812  
 VERSION BF526812.1 GI:11614175  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 1184)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 COMMENT  
 Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rtmail.nih.gov  
 Tissue Procurement: David N. Louis, M.D.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: L1AM9785 row: 0 column: 16  
 High quality sequence stop: 565.  
 Location/Qualifiers

FEATURES  
 source  
 1..1184  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4213479"  
 /clone\_lib="NCI CGAP\_Brn64"  
 /tissue\_type="glio/blastoma with EGFR amplification"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /note="Organ: Brain; Vector: PCMV-SPORE; Site:1: NotI;  
 Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT.  
 Average insert size 1.57 kb. Constructed by Life  
 Technologies. Note: this is a NCI CGAP Library."

BASE COUNT 435 a 235 c 269 g 245 t

Alignment Scores:

Pred. No.: 0.0126 Length: 1184  
 Score: 13.00 Matches: 13  
 Percent Similarity: 100.00% Conserved: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 13.40% Indels: 0  
 DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BF526812 (1-1184)

QY 8 HisAlaSerLeuGlyAspSerGlyThrLeuSerGlyThr 20  
 Db 61 CACGCTAGCCTGGGTGACAGTGAACGCTGTCTCAACA 99

RESULT 8  
 AZ020264/c

LOCUS 357 bp DNA linear GSS 25-FEB-2000  
 DEFINITION RPCI-23-299J8.TV RPCI-23 Mus musculus genomic clone RPCI-23-299J8,  
 DNA sequence.  
 ACCESSION AZ020264  
 VERSION AZ020264.1 GI:7095648  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

REFERENCE  
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 357)  
Zhao, S., Nierman, W., Felblyum, T., Malek, J., Shatsman, S., Akiret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C.M.

TITLE  
JOURNAL  
COMMENT

Mouse BAC End Sequences from Library RPCI-23  
Unpublished (1999)  
Other\_GSSs: RPCI-23-299J8.TJ

Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieder@jeng.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>) or from Reseach Genetics (info@resgen.com). BAC end page: [http://www.tigr.org/tdb/bac/ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac/ends/mouse/bac_end_intro.html)  
Plate: 299 row: J column: 8  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source

Location/Qualifiers  
1..357  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-23-299J8"  
/clone\_11b="RPCI-23"  
/sex="Female"  
/lab\_host="DH10B"  
/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site: 1; EcoRI; Site: 2; EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
BASE COUNT 49 a 82 c 27 g 194 t 5 others  
ORIGIN

## Alignment Scores:

Pred. No.: 0.048 Length: 357  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
Gaps: 17

US-09-854-133-586 (1-97) x AZ020264 (1-357)

OY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||  
Db 245 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 210

RESULT 9  
LOCUS AZ799582 589 bp DNA linear GSS 16-FEB-2001  
DEFINITION 2M0057F09F Mouse 10kb plasmid U0GC1M library Mus musculus genomic  
clone U0GC2M0057F09 F, DNA sequence.

ACCESSION AZ799582  
VERSION AZ799582.1 GI:12950846  
KEYWORDS GSS  
SOURCE mouse mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 589)  
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL  
COMMENT

Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu

Insert Length: 10000 Std Error: 0.00  
Plate: 0057 row: F column: 09  
Seq primer: CGTTGTAAACGACGCCACGT  
Class: plasmid ends  
High quality sequence stop: 589.

FEATURES  
source

Location/Qualifiers  
1..589  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U0GC2M0057F09"  
/clone\_11b="Mouse 10kb plasmid U0GC1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: pMD42uv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/narres/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (9147321419b1AF19072.1), a copy number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 80 a 170 c 91 g 248 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0.0709 Length: 589  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 12.37% Indels: 0  
Gaps: 17

US-09-854-133-586 (1-97) x AZ799582 (1-589)

OY 23 ArgLysLysGluArgLysLysLysArgGluArgLys 34  
|||||  
Db 168 AGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 133

RESULT 10  
LOCUS AQ351658 629 bp DNA linear GSS 24-JAN-1999  
DEFINITION CITR1-EI-2529K22.TF CITR1-EI Homo sapiens genomic clone 2529K22,  
DNA sequence.

ACCESSION AQ351658  
VERSION AQ351658.1 GI:4178993  
KEYWORDS GSS  
SOURCE human.  
ORGANISM Homo sapiens

## REFERENCE AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 629)  
Zhao, S. Adams, W. D.

**TITLE**

JOURNAL  
COMMENT

Venter, J.C., Karp, P., Kohn, M., Lander, E.S., Smith, S.M., White, O., Adams, M.D., Holt, R.A., Peterson, J., Shizuya, H., Simon, M., and Unger, K. (1997) *Genomic Libraries for Sequence-Ready Mapping*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

## FEATURES

Email: [hbe@tigr.org](mailto:hbe@tigr.org)  
 Clones are available from Research Genetics ([info@resgen.com](mailto:info@resgen.com)). BAC  
 end search page:  
[http://www.tigr.org/ttd/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/ttd/humgen/bac_end_search/bac_end_search.html).  
 Seq primer: M13-21  
 Class: BAC ends.

son

Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeetigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page: [http://www.tigr.org/cdb/humgen/Bac\\_end\\_search/Bac\\_end\\_search.html](http://www.tigr.org/cdb/humgen/Bac_end_search/Bac_end_search.html).  
Seq primer: M13-21  
Class: BAC ends.

FEATURES	Location/Qualifiers
source	1. .629

BASE COUNT	112 a	170 c	96 g	251 t
ORIGIN	Caltech Human BAC Library D			

### Alignment Scores:

Pred. No.:	0.0746	Length:	62
Score:	12.00	Matches:	12
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	12.37%	Indels:	0
DB:	17	Gaps:	0

US-09-854-133-586 (1-97) x AQ351658 (1-629)

QY	23	ArgIysLysGIuArgLysLysAArgLIuArgLys	34
Db	502	AGAAAGAAAGAAAGAAAAAGAGAAAGAAAG	46

RESULT 11  
AQ352329/c

LOCUS	DEFINITION	ACCESSION
AO352329	629 bp DNA	linear
CITBI-EI-2529M24.TF	CITBI-EI Homo sapiens genomic clone	GSS 24-JAN-1999
AO352329	DNA sequence.	2529M24'

VERSION  
KEYWORDS  
GSS.  
AQ352329.1  
GI:4179664

## ORGANISM

**REFERENCE**

Zhao S, Zhang J, Wang Y, et al. 2019. The phylogenetic relationships among Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

TITLE	Author
Venter, J.C.	Venter, J.C., Malek, J., Shlezuya, H., Simon, M. and
Use of BAC End Sequences from Caltech Ribosomal	

JOURNAL  
COMMENT

Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208

BASE COUNT	ORIGIN	Caricoll Human BAC Library D <sup>a</sup>
112 a	169 c	96 g 252 t

Alignment Scores:  
Pred. No. :  
Score.

Score:	12.00	Length:	62
Percent Similarity:	100.00%	Matches:	12
Best Local Similarity:	100.00%	Conservative:	0
Query Match:	12.37%	Mismatches:	0
DB:	17	Indels:	0

05-834-133-586 (1-97) x AQ352329 (1-629)

QY	23	ArgLysIysSGLuArgIysLysLysaRgLuArgIys	34
Db	502	AGAAAGAAAGAAAGAAAGAAAGAGAAAGAAAG	46

RESULT 12  
AZ258055/  
LOCUS

DEFINITION	664 bp	DNA	linear	GSS 26-JUL-2000
RPCT-23-134E11.TV	RPCT-23	Mus musculus	genomic clone	RPCT-23-134E11
ACCESSION	/ DNA sequence.			
DESCRIPTION	/			

VERSION AZ25805.1 GI:9463244  
KEYWORDS GSS.

## ORGANISMS

REFERENCE AUTHORS	TITLE JOURNAL	COMMENT
Eukaryota: Mammalia: Euteleostomi: 1 (bases 1 to 664)	Chordata: Craniata: Vertebrata: Euteleostomi: 1 (bases 1 to 664)	
Zhao S, Nierman, W., Feldblum, T., Malek, J., Shatsman, S., Akhmet , B., Levins, M., McGinn, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P and Fraser, C. N.	Mouse BAC End Sequences from Library RPCI-23 unpublished (1999)	
Other GSS: RPCI-23-134E11.TJ		

Department of Eukaryotic Genomics  
The Institute for Genomic Research

Tel: 301 838 0200  
Fax: 301 838 0208

Email: [szhao@tigr.org](mailto:szhao@tigr.org)  
Clones are derived from  
library available at:

information received from the mouse BAC library RPCR-23. For BAC library availability, please contact Plietser de Jong (plieter@edjong.med.buffalo.edu). Clones may be purchased from BACpac Resources (<http://bacpac.med.buffalo.edu/orderingframe.html>) or from Resea ch Genetics (info@resgen.com). BAC end page: [http://www.tigr.org/tdb/bac\\_ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html)  
 Plate: 134 row: E column: 11  
 Seq primer: 77  
 Class: BAC ends.

FEATURES  
source

```
Location/Qualifiers
1. .664
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
```



	library was cloned in the pARMBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."
BASE COUNT	198 a 174 c 77 g 422 t
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0961 Length: 871
Score:	12.00 Matches: 12
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mistmatches: 0
Query Match:	12.37% Indels: 0
DB:	17 Gaps: 0
US-09-854-133-586 (1-97) x A2747681 (1-871)	
QY 23 ArgLysLySGluArgLysLySLysArgGluArgLys 34	
Dn 102 AGAAGCAAAAGAAAAGAAGAAAAAGAGAGAGAAAA 67	
RESULT 14 BE880444	
LOCUS BE880444	882 bp mRNA linear EST 20-OCT-2000
DEFINITION 601491430F1 NIH_MGC_69 Homo sapiens CDNA clone IMAGE:3893886 5'	
ACCSSION BE880444	mRNA sequence.
VERSION BE880444	BE880444 GI:10329220
KEYWORDS EST.	
SOURCE human.	
ORGANISM Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominiade; Homo.	
REFERENCE 1 (bases 1 to 882)	
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.	
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL Unpublished (1999)	
COMMENT Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DTP/Gazdar CDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/JLNt info: <a href="http://image.llnl.gov">http://image.llnl.gov</a> Plate: LHAM9682 row: k column: 07 High quality sequence stop: 662. Location/Qualifiers 1..882 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:3893886" /clone_lib="NIH_MGC_69" /tissue_type="large cell carcinoma, undifferentiated" /lab_host="DH10B (phage-resistant)" /note="Organ: lung; Vector: PCMV-SPOrt6; Site_1: NotI; site_2: SalI; Cloned unidirectionally. Primer: Oligo dtf. Average insert size 1.1 kb. Library constructed by Life Technologies"	
FEATURES	
source	
BASE COUNT	263 a 165 c 182 g 272 t
ORIGIN	
Alignment Scores:	
Pred. No.:	0.0971 Length: 882
Score:	12.00 Matches: 12
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mistatches: 0
Query Match:	12.37% Indels: 0
DB:	12 Gaps: 0
US-09-854-133-586 (1-97) x BE880444 (1-882)	
QY 8 HisAlaserLeucIGlyasperGIutIRheSerCln 19	

```

Db      14  CACGCTAGCCTGGTGACAGTGCAGCCCTGCTCA 49
RESULT  15
LOCUS   BM557601
DEFINITION  BM557601 1032 bp mRNA linear EST 20-FEB-2002
5', mRNA sequence.
ACCESSION  BM557601
VERSION    BM557601.1 GI:18799717
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  NIH-MGC http://mgi.nci.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            Unpublished (1999)
AUTHORS    Contact: Robert Strausberg, Ph.D.
            Email: cgabbs-r@mail.nih.gov
JOURNAL    Tissue Procurement: ATCC
COMMENT    cDNA Library Preparation: Life Technologies, Inc.
            DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LLM12754 row: c column: 20
            High quality sequence stop: 588.
FEATURES
SOURCE
1. 1032
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5739979"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dr primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH-MGC Library."
BASE COUNT 301 a 196 c 214 g 320 t
ORIGIN
Alignment Scores:
Pred. No.: 0.11 Length: 1032
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.37% Indels: 0
DB: 13 Gaps: 0
US-09-854-133-586 (1-97) x BM557601 (1-1032)
QY 8 H1A1aserleugiyaspserclutrlleusergln 19
Db 95 CACGCTAGCCTGGTGACAGTGCAGCCCTGCTCA 130
RESULT 16
LOCUS   BF809298 210 bp mRNA linear EST 12-JAN-2001
DEFINITION  BF809298
ACCESSION  BF809298
VERSION    BF809298.1 GI:12138287
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 210)
            1
AUTHORS

```

```

Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL    MEDLINE
MEDLINE    20202663
COMMENT    Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704932
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?l=QV0&t2=QV0-CI0196-
            151100-505-a08&t3=2000-11-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 8
            High quality sequence stop: 210.
FEATURES
SOURCE
1. 210
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CI0196"
/dev_stage="Adult"
/note="Organ: colon; ins; Vector: puc18; site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 29 a 50 c 48 g 83 t
ORIGIN
Alignment Scores:
Pred. No.: 0.307 Length: 210
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.34% Indels: 0
DB: 12 Gaps: 0
US-09-854-133-586 (1-97) x BF809298 (1-210)
QY 23 ArgLysLysGluArgLysLysLysArgGluArg 33
Db 115 AGAAAGAGGAGAAAGAAAGAAAGAGAGAGAGA 83
RESULT 17
LOCUS   BF743959 227 bp mRNA linear EST 10-JAN-2001
DEFINITION  BF743959
ACCESSION  BF743959
VERSION    BF743959.1 GI:12070635
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE  Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 227)
            1
AUTHORS    Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,
            Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare
            M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed

```

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE  
20202663  
COMMENT  
Contact: Simpson A.J.G.

BASE COUNT ORIGIN	36 a	64 c	37 g	90 t
----------------------	------	------	------	------

US-09-854-133-586 (1-97) x BF743959 (1-227)

## FEATURES

Location/Qualifiers  
1. .259

/organism="Mus musculus"  
/strain="C57BL/6j"  
/db\_xref="taxon:10090"  
/clone="Rpci-23-110p5"  
/clone\_lib="Rpci-23"  
/sex="Female"  
/lab\_host="DH10B"

/note="Organ: Kidney/Brain; Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; Female C57BL/6j mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."  
BASE COUNT 11 a 90 c 27 g 131 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0.362 Length: 259  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x A2278638 (1-259)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
DB 218 AGAAGAAAGAAAGAAAGAAAGAAAGAGAGAGA 186

## RESULT 20

BO558860/c

LOCUS BO558860 374 bp mRNA linear EST 20-JUN-2002  
DEFINITION H4054G03-3 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone  
ACCESSION H4054G03.3 mRNA sequence.  
VERSION BO558860  
KEYWORDS BO558860.1 GI:21459745  
SOURCE EST

ORGANISM house mouse.  
Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
AUTHORS 1 (bases 1 to 374)  
VanBuren V., Piao Y., Dudekula D.B., Qian Y., Carter M.G., Martin P.R., Stagg C.A., Bassey V., Alba K., Hamatani T., Kargul G.J., Luo A.G. and Ko M.S.H.

TITLE Assembly, verification, and initial annotation of NIA 7.4K mouse cDNA clone set  
JOURNAL Unpublished (2002)  
COMMENT Other ESTs: H4054G03-5  
Contact: Yong Qian  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA  
Email: cdmagel@nigms.nih.gov  
This clone set has been freely distributed to the community. Please visit [http://lgsun.grc.nia.nih.gov/cDNA/NIA\\_7.4K.html](http://lgsun.grc.nia.nih.gov/cDNA/NIA_7.4K.html) for details.  
Plate: H4054 row: G column: 03  
Seq primer: -21M13 Forward  
High quality sequence stop: 374  
POLYA=Yes.

## FEATURES

Location/Qualifiers  
1. .374

/organism="Mus musculus"  
/strain="C57BL/6j"  
/db\_xref="taxon:H4054G03-3"  
/db\_xref="taxon:10090"  
/clone="H4054G03"  
/clone\_lib="NIA Mouse 7.4K cDNA Clone Set"  
/sex="mixed"

/dev\_stage="mixed"  
/lab\_host="DH10B"  
/note="Vector: pSPOR1; Site\_1: SalI; Site\_2: NotI; This clone is among a rearranged set of 7,407 clones from more than 20 cDNA libraries."  
BASE COUNT 56 a 99 c 48 g 171 t  
ORIGIN

## Alignment Scores:

Pred. No.: 0.481 Length: 374  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 14 Gaps: 0

US-09-854-133-586 (1-97) x BO558860 (1-374)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
DB 64 AGAAGAAAGAAAGAAAGAAAGAAAGAGAGAGA 32

## RESULT 21

AI628395/c

LOCUS AI628395 382 bp mRNA linear EST 07-MAR-2000  
DEFINITION ly76a03.x1 NCI-CGAP Kid1 Homo sapiens cDNA clone IMAGE:2284972 3' similar to contains Alu repetitive element; contains LI.B1 LI repetitive element ;, mRNA sequence.  
ACCESSION AI628395  
VERSION AI628395.1 GI:4665195  
KEYWORDS EST  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
AUTHORS 1 (bases 1 to 382)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov  
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: [www.bio.lnl.gov/biopr/image/image.html](http://www.bio.lnl.gov/biopr/image/image.html)  
Insert length: 777 Std Error: 0.00  
Seq primer: -40UP from Glibco  
High quality sequence stop: 338  
POLYA=No.

## FEATURES

Location/Qualifiers  
1. .382

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2284972"  
/clone\_lib="NCI-CGAP\_Kid1"  
/lab\_host="DH10B"  
/note="Organ: Kidney; Vector: pRTT3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP hybridization, this DNA was used as tracer in a subtractive from a pool of 5,000 clones made from the same library (clones 132376-132391, 1456007-1456775, and 1500552-1500855). Subtraction by Bento Soares and M. Fatima Bonaldo."  
BASE COUNT 109 a 60 c 68 g 145 t  
ORIGIN

## Alignment Scores:

Pred. No.:	0.489	Length:	382
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.34%	Indels:	0
DB:	9	Gaps:	0

US-09-854-133-586 (1-97) x A1628395 (1-382)

QY 10 Serleuglyaspsergluthtleuserglunhr 20  
 |||||  
 DB 318 AGCCTGGCGACAGTGTCTCTCAACCA 286

## RESULT 22

BM196350/c 442 bp mRNA linear EST 30-JAN-2002  
 LOCUS C0327E04-3 NIA Mouse Undifferentiated ES Cell cDNA Library (long)  
 Mus musculus cDNA clone C0327E04 3', mRNA sequence.

ACCESSION BM196350  
 VERSION BM196350.1 GI:17748125

KEYWORDS EST.  
 SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 442)

REFERENCE Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Lim, M.K., Luo, A.,  
 Jaradat, S.A., Boheler, K.R. and Ko, M.S.H.  
 Systematic Analyses of NIA Mouse Undifferentiated ES Cell cDNA  
 Library (long)

## JOURNAL COMMENT

Unpublished (2001)  
 Laboratory of Genetics  
 National Institute on Aging/National Institutes of Health  
 333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
 Email: cdna@igsun.grc.nia.nih.gov  
 Plate: C0327 row: E column: 04  
 Seq primer: -21M13 Forward  
 High quality sequence stop: 442  
 POLYA=Yes.

## FEATURES

Location/Qualifiers

1..442  
 /organism="Mus musculus"  
 /strain="129/Sv x 129/Sv-CP"  
 /db\_xref="taxon:C0327E04-3"  
 /db\_xref="taxon:10090"  
 /clone="C0327E04"  
 /clone\_lib="NIA Mouse Undifferentiated ES Cell cDNA  
 Library (long)"  
 /tissue\_type="Undifferentiated ES Cell"  
 /cell\_line="R1 ES cells"  
 /lab\_host="DHI0B"  
 /note="Vector: pSPORT1 (Invitrogen); Site.1: SalI; Site.2:  
 NotI; Mouse cDNA project by the Laboratory of Genetics,  
 National Institute on Aging (NIA), Intramural Research  
 Program, NIH (http://igsun.grc.nia.nih.gov/cDNA). This is a  
 long-transcript enriched cDNA library (Ref. Genome Res.  
 11: 1553-1558 (2001)). (PMID: 11344193). Total RNAs were  
 obtained from Dr. Kenneth R. Boheler (National Institute  
 on Aging, USA). ES cells were cultured without feeder  
 cells in the presence of LIF and BRL-conditioned media.  
 Double-stranded cDNAs were synthesized with an Oligo(dT)  
 primer (Invitrogen):  
 5'-pGACTAGTCTAGATCGGACGCGCCCTTTTCTTTT-3' from  
 14.2 ug of total RNA, treated with T4 DNA polymerase, and  
 purified by ethanol-precipitation. The cDNAs were ligated  
 to lone-linker LP-SalI, purified by phenol/chloroform, and  
 separated from free linkers by Centricon 100. Then, the  
 cDNAs were amplified by long-range high fidelity PCR using  
 Ex Taq polymerase (Takara) with a primer SalI-S. The  
 products were purified by phenol/chloroform and Centricon

## BASE COUNT

100. The cDNAs were digested with SalI and NotI enzymes  
 and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
 The DH10B E. coli host was transformed with the ligation  
 mixture by the standard chemical method. The average  
 insert size is about 2.4 kb. The library was constructed  
 by Yulan Piao (NIA)."

## Alignment Scores:

Pred. No.:	0.548	Length:	442
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	11.34%	Indels:	0
DB:	13	Gaps:	0

US-09-854-133-586 (1-97) x BM196350 (1-442)

QY 23 Arglyslgylargylslgylslyargyluarg 33  
 |||||  
 DB 64 AGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 32

## RESULT 23

BH117011/c 445 bp DNA linear GSS 19-JUL-2001  
 LOCUS RPCI-24-358K21.TJ RPCI-24 Mus musculus genomic clone RPCI-24-358K21  
 , DNA sequence.

ACCESSION BH117011  
 VERSION BH117011.1 GI:14957851

KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 445)

REFERENCE Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akiref, B., Lewins, M.,  
 Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,  
 Russell, D., de Jong, P., and Fraser, C.M.  
 Mouse BAC End Sequences from Library RPCI-24

Unpublished (1999)  
 Other\_GSSs: RPCI-24-358K21.TJ

CONTACT: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208

Email: szhao@tigr.org  
 Clones are derived from the mouse BAC library RPCI-24. For BAC  
 library availability, please contact Pieter de Jong  
 (pdejong@mail.cho.org). Clones may be purchased from BACPAC  
 Resources (http://www.choi.org/bacpac/orderingframe.htm). BAC end  
 page: http://www.tigr.org/tdb/bac\_ends/mouse/bac\_end\_intro.html  
 Plate: 358 row: K column: 21  
 Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1..445  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="RPCI-24-358K21"  
 /clone\_lib="RPCI-24"  
 /sex="Male"  
 /cell\_type="Spleen/Brain"  
 /note="Vector: pTARBAC1; Site.1: BamHI; Site.2: BamHI;  
 RPCI-24 Mouse BAC Library produced by Pieter de Jong. The  
 library was cloned in the pTARBAC1 cloning vector at the  
 BamHI sites using MboI partially digested male C57BL/6J  
 DNA."

## FEATURES

Location/Qualifiers

BASE COUNT 70 a 102 c 81 g 192 t  
 ORIGIN

1

Early Match:	11.348	Indels:
	17	Gaps:

• • • • •

US-09-854-133-586 (1-97) x AZ356753 (1-464)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
 DB 405 AGAAGAAAGAAAGAAAGAAAGAGAGAGA 437

## RESULT 26

AZ779079 495 bp DNA linear GSS 16-FEB-2001  
 LOCUS 2M0014P19R Mouse 10kb plasmid U0GC1M library Mus musculus genomic  
 DEFINITION clone U0GC2M0014P19 R, DNA sequence.

ACCESSION AZ779079  
 VERSION AZ779079.1 GI:12909373

KEYWORDS GSS.

SOURCE house mouse.

## ORGANISM

Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 495)  
 Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,  
 Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,  
 M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.  
 and Wright, D., Weiss, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert length: 1000 Std Error: 0.00

Plate: 0014 Row: P Column: 19

Seq primer: CACACAGGAAACAGCATATACAC

Class: plasmid ends

High quality sequence stop: 495.

Location/Qualifiers

1..495

/organism="Mus musculus"

/strain="C57BL/6J"

/db\_xref="taxon:10090"

/clone="U0GC2M0014P19"

/clone\_lib="Mouse 10kb plasmid U0GC1M library"

/sex="Male"

/lab\_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMDA2nv. Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD42 (g11473114[gb|AF129072.1]), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance.

BASE COUNT 107 a 130 c 52 g 206 t

## ORIGIN

Alignment Scores: 0.599 Length: 495  
 Pred. No.: 0.599

Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.34% Indels: 0  
 DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AZ779079 (1-495)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
 DB 90 AGAAGAAAGAAAGAAAGAAAGAGAGAGA 58

## RESULT 27

BH858364 500 bp DNA linear GSS 08-JUL-2002  
 LOCUS B5\_164a2.s6 Mouse Retroviral Tagged Cancer Gene Database Mus  
 DEFINITION musculus genomic clone B5\_164a2, DNA sequence.

ACCESSION BH858364  
 VERSION BH858364.1 GI:21709185

## KEYWORDS

GSS.

## SOURCE

house mouse.  
 Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 500)  
 Jenkins, N.A. and Copeland, N.G.  
 Retroviral tagging provides a potent cancer gene discovery tool in  
 the post-genome-sequence era

Nat. Genet., (2002) In press

Contact: Copeland NG

Mouse Cancer Genetics Program

National Cancer Institute

Bldg. 539, Rm. 229, Frederick, MD 21702-1201, USA

Tel: 301 846 1260

Fax: 301 846 6666

Email: copeland@ncifcrf.gov

Class: PCR with specific primers.

Location/Qualifiers

1..500

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="B5\_164a2"

/clone\_lib="Mouse Retroviral Tagged Cancer Gene Database"

/sex="female"

/tissue\_type="leukemia"

/note="Inverse PCR method

(http://genome2.ncifcrf.gov/RTCGD)"

BASE COUNT 101 a 111 c 106 g 182 t

## ORIGIN

## Alignment Scores:

Pred. No.: 0.604 Length: 500  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.34% Indels: 0  
 DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x BH858364 (1-500)

QY 24 LysLysGluArgLysLysArgGluArgLys 34  
 DB 233 AAAAAAGAAAGAAAGAAAGAGAGAGAGAAA 201

## RESULT 28

AZ750540 506 bp DNA linear GSS 25-JAN-2001  
 LOCUS AZ750540  
 DEFINITION RPCI-24-112119.TJ RPCI-24 Mus musculus genomic clone RPCI-24-112119  
 , DNA sequence.

ACCESSION AZ750540  
 VERSION AZ750540.1 GI:12535699  
 KEYWORDS GSS.

SOURCE house mouse.  
ORGANISM Mus musculus  
REFERENCE 1 (bases 1 to 506)  
AUTHORS Zhao, S., Nierman, W., Malek, J., Shatsman, S., Aktinret, B., Levins, M., Russell, D., de Jong, P., and Fraser, C.M.  
TITLE Mouse BAC End Sequences from Library RPCI-24  
JOURNAL Unpublished (1999)  
COMMENT Other GSSs: RPCI-24-112119.TV  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@tigr.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end plate: [http://www.tigr.org/tldb/bac-ends/mouse/bac\\_end\\_intro.html](http://www.tigr.org/tldb/bac-ends/mouse/bac_end_intro.html).  
Seq primer: SP6  
Class: BAC ends.

#### FEATURES

Location/Qualifiers  
1..506  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="RPCI-24-112119"  
/clone\_1lb="RPCI-24"  
/sex="Male"  
/cell\_type="Spleen/Brain"  
/note="Vector: pPARBAC1, Site\_1: BamHI, Site\_2: BamHI.  
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pPARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."  
BASE COUNT 232 a 69 c 131 g 74 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 0.609 Length: 506  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x A2750540 (1-506)

OY 23 ArgLysIysGluArgLysLysLysArgGluArg 33  
Db 130 AGAAGAAAGAAAGAAAGAAAGAAAGAGAGAGA 162

#### RESULT 29

LOCUS B0086880 530 bp mRNA linear EST 05-APR-2002  
DEFINITION ih91d05.y1 Melton Mouse E16 5 Pancreas Library 2 M16B2 Mus musculus  
ACCESSION B0086880  
VERSION B0086880.1 GI:20046084  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 530)  
AUTHORS Melton, D., Brown, D., Kenty, G., Permutt, A., Lee, C., Kaestner, K., Lemishka, I., Searce, M., Brestelli, J., Gradwohl, G., Clifton, S., Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Birstein, A.,

Schmitt, A., Theising, B., Ritter, E., Ronko, J., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagareishvili, R., Williams, T., Jackson, Y., and Bowers, Y.  
TITLE Endocrine Pancreas Consortium  
JOURNAL Unpublished (2000)  
COMMENT Other ESTs: ih91d05.x1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@biohpc.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by: Washington University Genome Sequencing Center This clone is available royalty-free through LBNL; please contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information  
Seq primer: -40RP from Gibco  
High quality sequence stop: 432.

#### FEATURES

Location/Qualifiers  
1..530  
/organism="Mus musculus"  
/strain="ICR"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5939385"  
/clone\_1lb="Melton Mouse E16 5 Pancreas Library 2 M16B2"  
/sex="Both"  
/tissue\_type="Total pancreas"  
/dev\_stage="Embryonic day 16.5"  
/lab\_host="TOP10"  
/note="Organ: Pancreas; Vector: pBluescript II SK; Site\_1: NotI; Site\_2: SalI; Library constructed using Superscript oligo-dT priming. Size-selected by column fractionation; average insert size 1.06kb. Primary library, unamplified."  
BASE COUNT 244 a 74 c 116 g 96 t  
ORIGIN

#### Alignment Scores:

Pred. No.: 0.632 Length: 530  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 14 Gaps: 0

US-09-854-133-586 (1-97) x B0086880 (1-530)

OY 23 ArgLysIysGluArgLysLysLysArgGluArg 33  
Db 267 AGAAGAAAGAAAGAAAGAAAGAAAGAGAGAGA 299

#### RESULT 30

LOCUS A0547131/C 583 bp DNA linear GSS 28-MAY-1999  
DEFINITION RPCI-11-431G8.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-431G8, A0547131  
ACCESSION A0547131  
VERSION A0547131.1 GI:4906374  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 583)  
AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P., and Venter, J.C.  
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready Map Building  
JOURNAL Unpublished (1997)  
COMMENT Other GSSs: RPCI-11-431G8.TV



Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeet@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@edj.med.buffalo.edu). Clones may be purchased from BACRAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from Research Genet cs (<http://inforesgen.com>). BAC end search page: [http://www.tigr.org/tdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html).  
Seq primer: SP6  
Class: BAC ends.

#### FEATURES

Location/Qualifiers  
1..583  
/organism="Homo sapiens"  
/db\_xref="GDB:763271"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-431G8"  
/clone\_1lb="RPCI-11"  
/sex="Male"  
/cell\_type="lymphocytes"  
/note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI; RPCI11 Human Male BAC library"

BASE COUNT 81 a 130 c 50 g 322 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 0.68 Length: 583  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AG547131 (1-583)

QY 24 LysLysGluArgLysLysArgGluArgLys 34  
Db 274 AAGAAAGAAAGAAAGAAAGAAAGAAAG 242

#### RESULT 31

LOCUS B0086621 587 bp mRNA linear EST 05-APR-2002  
DEFINITION B0086621  
CDNA clone IMAGE:5939385 3', mRNA sequence.

ACCESSION B0086621  
VERSION B0086621.1 GI:20045825  
KEYWORDS EST.  
SOURCE house mouse.  
ORGANISM Mus musculus

REFERENCE  
AUTHORS Mus musculus; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 587)

TITLE  
JOURNAL Lemmon, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K., Hillier, L., Marra, M., Page, D., Wylie, T., Martin, J., Hristova, A., Schmitt, A., Theising, B., Ritter, E., Konko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, T., Jackson, Y., and Bowers, Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@molb.harvard.edu  
Library was constructed by Dr. Douglas Melton DNA sequencing by:

Washington University Genome Sequencing Center This clone is available royalty-free through LBNL; please contact the IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information  
Seq primer: -40UP from Glibco  
High quality sequence stop: 466.  
Location/Qualifiers  
1..587  
/organism="Mus musculus"  
/strain="ICR"  
/db\_xref="taxon:10090"  
/clone="IMAGE:5939385"  
/clone\_1lb="Melton Mouse E16 5 Pancreas Library 2 M1B2"  
/sex="Both"  
/tissue\_type="Total pancreas"  
/dev\_stage="Embryonic day 16.5"  
/lab\_host="TOP10"  
/note="Organ: Pancreas; Vector: pBluescript II SK; Site\_1: NotI; Site\_2: SalI; Library constructed using SuperScript Plasmid Library kit (Life Technologies). cDNA made by oligo-dr priming. Size-selected by column fractionation; average insert size 1.06kb. Primary library, unamplified."

BASE COUNT 170 a 106 c 88 g 223 t  
ORIGIN

Alignment Scores:  
Pred. No.: 0.684 Length: 587  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 14 Gaps: 0

US-09-854-133-586 (1-97) x B0086621 (1-587)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
Db 564 AAGAAAGAAAGAAAGAAAGAAAGAAAG 532

#### RESULT 32

LOCUS AG075749 643 bp DNA linear GSS 03-NOV-2001  
DEFINITION Pan troglodytes DNA, clone: PTB-068123.F, genomic survey sequence.  
ACCESSION AG075749  
VERSION AG075749.1 GI:16627551  
KEYWORDS GSS.  
SOURCE Pan troglodytes male lymphoblast DNA, clone\_1lb:PTB Chimpanzee Male BAC library clone:PTB-068123.F.

ORGANISM Pan troglodytes  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
BAC end sequences of Library PTB

TITLE  
JOURNAL Unpublished  
2 (bases 1 to 643)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: [chimpsesgsc.riken.go.jp](mailto:chimpsesgsc.riken.go.jp), URL: <http://hgp.gsc.riken.go.jp/>, Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.

PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pKS145  
R.Site 1 : SacI

FEATURES  
Source  
R.Site 2 : Saci.  
Location/Qualifiers  
1.643  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-068123.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_lib="PTB Chimpanzee Male BAC Library"  
1 others

BASE COUNT 165 a 144 c 128 g 205 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.734 Length: 643  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AG075749 (1-643)

OY 24 LysLysGUAArgLysLysArgGUAArgLys 34  
Db 566 AAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 534

RESULT 33  
A2791667 652 bp DNA linear GSS 16-FEB-2001  
LOCUS 2M0041K09 Mouse 10kb plasmid UUCGCM library Mus musculus genomic  
ACCESSION A2791667  
VERSION A2791667.1 GI:12934796  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 652)  
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,  
Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rilly  
and Wright, D., Rose, R., Stokes, R., Tinney, A., von Niederhausern, A.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0041 row: K column: 09  
Seq primer: CACACAGGAAACACGTATGACC  
Class: plasmid ends  
High quality sequence stop: 652.  
Location/Qualifiers  
1.652  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG2M0041K09"  
/clone\_lib="Mouse 10kb plasmid UUCGCM library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv; Purified genomic DNA from M.  
musculus C57BL/6J (male) was obtained from the Jackson  
Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA  
was hydrodynamically sheared by repeated passage through a.

0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g114732114[9b]AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 290 a 94 c 121 g 146 t 1 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.742 Length: 652  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatave: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 11.34% Indels: 0  
DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x A2791667 (1-652)

OY 23 ArgLysLysGUAArgLysLysArgGUAArg 33  
Db 368 AGAAAGAAAGAAAGAAAGAAAGAAAGAGAGAGA 400

RESULT 34  
AG126279/c 677 bp DNA linear GSS 04-NOV-2001  
LOCUS Pan troglodytes DNA, clone: PTB-136M16.F, genomic survey sequence.  
ACCESSION AG126279  
VERSION AG126279.1 GI:1655444  
KEYWORDS GSS.  
SOURCE Pan troglodytes male lymphoblast DNA, clone\_lib:PTB Chimpanzee Male  
BAC library clone:PTB-136M16.F.  
Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.  
1  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
BAC end sequences of library PTB  
Unpublished  
2 (bases 1 to 677)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (02-NOV-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Suihio-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:chiimpesegsc.riken.go.jp, URL:http://hsp.gsc.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB this BAC end  
was generated during the Rad process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13  
LIBRARY  
Vector : pRS145  
R.Site 1 : Saci  
R.Site 2 : Saci.  
Location/Qualifiers  
1.677  
/organism="Pan troglodytes"  
/db\_xref="taxon:9598"  
/clone="PTB-136M16.F"  
/sex="male"  
/cell\_type="lymphoblast"

```

BASE COUNT      169 a      220 c      153 g      135 t
ORIGIN
Alignment Scores:
Pred. No.:      0.764      Length:      677
Score:          11.00      Matches:      11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    11.34% Indels: 0
DB:             17 Gaps: 0

US-09-854-133-586 (1-97) x AG126279 (1-677)

QY      10 SerLeuGIYAspSerGIuThrLeuSerGIuThr 20
Db      416 AGCCTGGGGACGACGAGACCTGTCTCAACA 384

RESULT 35
LOCUS    A0947014      689 bp      DNA      linear      GSS 27-JAN-2000
DEFINITION Sheared DNA-45F24.TF Sheared DNA Trypanosoma brucei genomic clone
ACCESSION A0947014
VERSION   A0947014.1 GI:6770279
KEYWORDS  GSS.
SOURCE    Trypanosoma brucei.
ORGANISM  Trypanosoma brucei
           Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
REFERENCE 1 (bases 1 to 689)
AUTHORS   El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Sub,E., Malek,J., Fujii,C.,
           Gerrard,C., Leech,V., de Jong,P., Villu,E., Melville,S., Donelson,J.,
           Fraser,C. and Adams,M.
           Determination of clone end sequences from Trypanosoma brucei GUTat
           10.1 sheared DNA library
           Unpublished (1999)
           Other_GSSs: Sheared DNA-45F24.TF
           Contact: Najib M. El-Sayed
           Department of Eukaryotic Genomics
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: nelsayed@tigr.org
           Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
           DNA library constructed at TIGR. Clones will be available for
           distribution through ATCC. Sheared DNA end sequences search page:
           http://www.tigr.org/tdb/mbd/tbdb/.
           Seq primer: M13-Forward
           Class: Shotgun.
FEATURES             location/Qualifiers
source              1..689
                   /organism="Trypanosoma brucei"
                   /strain="TREU927/4 GUTat 10.1"
                   /db_xref="taxon:5691"
                   /clone="Sheared DNA-45F24"
                   /note="Vector: pUC18; Site_1: SmaI; Constructed at The
                   Institute for Genomic Research (TIGR), Rockville, MD.
                   Genomic DNA isolated from a cloned population of
                   Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                   sheared to give a fixed size distribution (approx 2 kb).
                   The v + l method used for the library construction is
                   described in detail in Smith, H.O. and Venter, J.C.
                   (Making small insert libraries for whole genome shotgun
                   sequencing projects. In Genome Sequencing: A Practical
                   Approach, eds. M. Vaubin and B. Bartell, Oxford University
                   Press, 1999)."
BASE COUNT      254 a      139 c      129 g      167 t
ORIGIN
Alignment Scores:

```

```

Pred. No.:      0.775      Length:      689
Score:          11.00      Matches:      11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    11.34% Indels: 0
DB:             17 Gaps: 0

US-09-854-133-586 (1-97) x A0947014 (1-689)

QY      23 ArgLYsLysGLuArgLYsLYsArgGLuArg 33
Db      94 AGAAAGAAAGAAAGAAAGAAAGAGAGAGA 126

RESULT 36
LOCUS    B1912987      696 bp      mRNA      linear      EST 16-OCT-2001
DEFINITION 603176462F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5240780 5',
ACCESSION B1912987
VERSION   B1912987.1 GI:16177268
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgabbs-remail.nih.gov
           Tissue Procurement: Life Technologies, Inc.
           cDNA Library Preparation: Life Technologies, Inc.
           CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMND)
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LMND at:
           http://image.llnl.gov
           Plate: LAM11607 row: c column: 21
           High quality sequence start: 3
           High quality sequence stop: 632.
FEATURES             location/Qualifiers
source              1..696
                   /organism="Homo sapiens"
                   /db_xref="taxon:9606"
                   /clone="IMAGE:5240780"
                   /clone_1b="NIH_MGC_121"
                   /lab_host="DH10B"
                   /note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
                   Site_2: EcoRV (destroyed); RNA source anonymous pool of 3
                   fetal brains, female age 20 weeks, female age 24 weeks,
                   and male age 26 weeks. Library is oligo-dT primed and
                   directionally cloned (EcoRV site is destroyed upon
                   cloning). Average insert size 1.7 kb, insert size range
                   0.7-3.5 kb. Library is normalized and enriched for
                   full-length clones and was constructed by C. Gruber
                   (Invitrogen). Research Genetics tracking code 017. Note:
                   this is a NIH_MGC Library."
BASE COUNT      186 a      144 c      184 g      182 t
ORIGIN
Alignment Scores:
Pred. No.:      0.781      Length:      696
Score:          11.00      Matches:      11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    11.34% Indels: 0
DB:             13 Gaps: 0

US-09-854-133-586 (1-97) x B1912987 (1-696)

QY      10 SerLeuGIYAspSerGIuThrLeuSerGIuThr 20

```



Plate: L1AM11607 row: b column: 21  
 High quality sequence start: 6  
 High quality sequence stop: 158.  
 Location/Qualifiers

FEATURES  
 source  
 1..1036  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5240756"  
 /clone\_lib="NIH\_MGC\_121"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMV-Sport6; Site: 1: NotI; Site: 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."  
 BASE COUNT 244 a 234 c 259 g 294 t 5 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 1.06 Length: 1036  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.34% Indels: 0  
 DB: 13 Gaps: 0

US-09-854-133-586 (1-97) x B1912939 (1-1036)

QY 10 SerLeuGIYASpSerGIuThLeuSerGIuThr 20  
 |||||  
 Db 77 AGCCTGGCGACACTGAGACCCCTCTCAACA 109

RESULT 40  
 LOCUS BC010121/c 2909 bp mRNA linear HTC 12-JUL-2001  
 DEFINITION Homo sapiens, similar to hypothetical protein MGC5149, clone  
 IMAGE:3840062, mRNA.  
 ACCESSION BC010121 GI:14714431  
 VERSION BC010121.1  
 KEYWORDS HTC.  
 SOURCE Homo sapiens.  
 ORGANISM Homo sapiens.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 2909)  
 AUTHORS Strausberg, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (02-JUL-2001) National Institutes of Health, Mammalian  
 Gene Collection (MGC), Cancer Genomics Office, National Cancer  
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
 USA

REMARK  
 COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
 Tissue Procurement: ATCC  
 CDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nsl.nih.gov](mailto:nisc_mgc@nsl.nih.gov)  
 Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,  
 Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,  
 Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,  
 Lim, M., Maduro, O.L., Mastaglio, C., Mastrian, S.D., McCloskey, J.C.,  
 McDowell, J., Pearson, R., Snyder, B., Stantip, S., Thomas, P.J.,  
 Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,

Zhang, L.-H. and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found  
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 27 Row: 1 Column: 10  
 This clone was selected for full length sequencing because it  
 passed the following selection criteria: Similarity but not  
 identity to protein  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers

FEATURES  
 source  
 1..2909  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3840062"  
 /tissue\_type="Placenta, choriocarcinoma"  
 /clone\_lib="NIH\_MGC\_21"  
 /lab\_host="DH10B-R"  
 /note="Vector: pOT87"  
 BASE COUNT 803 a 630 c 662 g 814 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.38 Length: 2909  
 Score: 11.00 Matches: 11  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 11.34% Indels: 0  
 DB: 11 Gaps: 0

US-09-854-133-586 (1-97) x BC010121 (1-2909)

QY 23 ArgLysLysGluArgLysLysArgGluArg 33  
 |||||  
 Db 93 CGAAGAGAGAGAGAGAGAGAGAGAGAGA 61

RESULT 41  
 LOCUS AZ742823/c 101 bp DNA linear GSS 25-JAN-2001  
 DEFINITION RPCI-24-15867.TJ RPCI-24 Mus musculus genomic clone RPCI-24-15867,  
 DNA sequence.  
 ACCESSION AZ742823  
 VERSION AZ742823.1 GI:12522079  
 KEYWORDS GSS.  
 SOURCE house mouse.  
 ORGANISM Mus musculus.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 REFERENCE 1 (bases 1 to 101)  
 AUTHORS Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akint, B., Levins, M.,  
 Tsegaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E.,  
 Russell, D., de Jong, P., and Fraser, C.M.  
 TITLE Mouse BAC End Sequences from Library RPCI-24  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shaying Zhao  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: [szhao@tigr.org](mailto:szhao@tigr.org)  
 Clones are derived from the mouse BAC library RPCI-24. For BAC  
 library availability, please contact Pieter de Jong  
 ([pdjong@tigr.org](mailto:pdjong@tigr.org)). Clones may be purchased from BACPAC  
 Resources (<http://www.chori.org/bacpac/orderingframe.htm>). BAC end  
 page: [http://www.tigr.org/tdb/bac\\_ends/mouse/BAC\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/mouse/BAC_end_intro.html)  
 Plate: 158 row: g column: 7  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers

1..101  
 /organism="Mus musculus"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"

```

/clone="RPCI-24-15867"
/clone.lib="RPCI-24"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: pTARBAcl; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAcl cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."
BASE COUNT      10 a      37 c      1 g      53 t
ORIGIN
Alignment Scores:
Pred. No.:      1.68      Length:      101
Score:          10.00     Matches:      10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    10.31% Indels: 0
DB:             17      Gaps: 0

US-09-854-133-586 (1-97) x AZ742823 (1-101)

QY      24      LyslySGluArgLysLysArgGluArg 33
Db      74      AAGAAAGAAAGAAAGAAAGAGAGAGC 45

RESULT 42
BI060988      113 bp      mRNA      linear      EST 15-JUN-2001
LOCUS      IL3-UT0116-020201-467-F11_1 UT0116 Homo sapiens cDNA, mRNA
DEFINITION      sequence.
ACCESSION      BI060988
VERSION      BI060988.1 GI:14468515
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 113)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL3&ct2=IL3-UT0116-
020201-467-F11_1&ts=2001-02-02&tt=1)
Seq primer: puc 18 forward
High quality sequence stop: 113.
Location/Qualifiers
1..113
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="UT0116"
/dev_stage="Adult"
/note="Organ: uterus,tumor; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESSES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)

```

```

profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      65 a      21 c      16 g      11 t
ORIGIN
Alignment Scores:
Pred. No.:      1.83      Length:      113
Score:          10.00     Matches:      10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    10.31% Indels: 0
DB:             13      Gaps: 0

US-09-854-133-586 (1-97) x BI060988 (1-113)

QY      10      SerLeuGlyAspSerGluThrLeuSerGln 19
Db      22      AGCCTGGGTGACACTGAGACTGTGCACAA 51

RESULT 43
BF853221      127 bp      mRNA      linear      EST 16-JUN-2001
LOCUS      MR2-EN0091-191200-002-909_1 EN0091 Homo sapiens cDNA, mRNA
DEFINITION      sequence.
ACCESSION      BF853221
VERSION      BF853221.1 GI:12240965
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 127)
AUTHORS      Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
COMMENT      Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&ct2=MR2-EN0091-
191200-002-909_1&ts=2000-12-19&tt=1)
Seq primer: puc 18 forward
High quality sequence stop: 67.
Location/Qualifiers
1..127
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone.lib="EN0091"
/dev_stage="Adult"
/note="Organ: lung,normal; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESSES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT      38 a      29 c      30 g      30 t
ORIGIN

```

## Alignment Scores:

Pred. No.: 2.01 Length: 127  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 12 Gaps: 0

US-09-854-133-586 (1-97) x BF853221 (1-127)

Oy 10 SerleuGLYASperGluThrleuSerGln 19  
 Db 24 AGCTGGGTGATAGCAGACTCTGTCCCA 53

## RESULT 44

AO239603/c

LOCUS

DEFINITION AO239603 142 bp DNA linear GSS 30-SEP-1998  
 C1T-HSP-2387119.TF.1 C1T-HSP Homo sapiens genomic clone 2387119,  
 DNA sequence.

ACCESSION AO239603.1 GI:3671894

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human.  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 142)  
 Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
 Berry,K., Granger,D., Suh,E., White,C., Shizuya,H., Simon,M. and  
 Venter,J.C.  
 Use of a random human BAC End Sequence Database for Sequence-Ready  
 Map Building  
 Unpublished (1998)  
 Contact: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: mdadams@tigr.org  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
 http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
 Seq primer: M13-21  
 Class: BAC ends.

## FEATURES

source

Location/Qualifiers  
 1..142  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="2387119"  
 /clone\_1ib="C1T-HSP"  
 /sex="Male"  
 /cell\_type="Sperm"  
 /note="Vector: pBelOBAC11; site\_1: HindIII; site\_2:  
 HindIII"  
 BASE COUNT 29 a 30 c 32 g 51 t  
 ORIGIN

## Alignment Scores:

Pred. No.: 2.19 Length: 142  
 Score: 10.00 Matches: 10  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 10.31% Indels: 0  
 DB: 17 Gaps: 0

US-09-854-133-586 (1-97) x AO239603 (1-142)

Oy 10 SerleuGLYASperGluThrleuSerGln 19

Db 106 AGCTGGGTGATAGCAGACTCTGTCCCA 77

RESULT 45

BB122746/c  
 LOCUS BB122746 155 bp mRNA linear EST 27-JUN-2000  
 DEFINITION BB122746 RIKEN full-length enriched, adult male urinary bladder Mus  
 musculus cDNA clone 9530089K22 3', mRNA sequence.  
 BB122746  
 BB122746.1 GI:8775314  
 EST.  
 house mouse.  
 Mus musculus

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 1 (bases 1 to 155)

AUTHORS

Kono,H., Aizawa,K., Akahira,S., Akiyama,J., Arakawa,T., Carninci  
 P., Endo,T., Fukuda,S., Fukunishi,Y., Hara,A., Hayatsu,N.,  
 Hirozane,T., Hori,F., Ishii,Y., Ishikawa,J., Ishikawa,T., Itoh,M.,  
 Iizawa,M., Kadota,K., Kagawa,I., Kai,C., Kawai,J., Kikuchi,N.,  
 Kiyosawa,H., Kojima,Y., Kondo,S., Koya,S., Kurihara,C., Kusabe,M.,  
 Matsuyama,T., Miki,R., Mizuno,Y., Nakamura,M., Oda,H., Okazaki,Y.,  
 Ono,T., Owa,C., Saito,H., Sakai,C., Sato,K., Shibata,K., Shibata  
 Y., Shigemoto,T., Shinagawa,A., Shiraki,T., Sogabe,Y., Sugahara,Y.,  
 Suzuki,H., Suzuki,H., Tagawa,A., Takahashi,F., Tomimaga,N., Toya  
 T., Tsunoda,Y., Watabiki,A., Watanabe,S., Yamamura,T., Yamane,I.,  
 Yano,R., Yasunishi,A., Yokota,T., Yoshida,K., Yoshiki,A., Yoshino  
 M., Muramatsu,M. and Hayashizaki,Y.  
 RIKEN Mouse ESTs (Kono,H., et al.)  
 Unpublished (2000)  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center(GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-resgsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Carninci,P., Nishiyama,Y., Westover,A., Itoh,M., Nagaoke,S., Sasaki  
 N., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
 Thermostabilization and thermoactivation of thermolabile enzymes by  
 trehalose and its application for the synthesis of full length  
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)  
 Itoh,M., Katsunai,T., Akiyama,J., Shibata,K., Iizawa,M., Kawai,J.,  
 Tomaru,Y., Carninci,P., Shibata,Y., Ozawa,Y., Muramatsu,M., Okazaki  
 Y. and Hayashizaki,Y.  
 Automated filtration-based high-throughput plasmid preparation  
 system. Genome Res. 9 (5), 463-470 (1999)  
 Carninci,P. and Hayashizaki,Y.  
 High efficiency full-length cDNA cloning. Methods Enzymol. 303,  
 19-44 (1999)  
 Please visit our web site (http://genome.rtc.riken.go.jp) for  
 further details.

## FEATURES

source

Location/Qualifiers  
 1..155  
 /organism="Mus musculus"  
 /db\_xref="taxon:10090"  
 /clone="9530089K22"  
 /clone\_1ib="RIKEN full-length enriched, adult male urinary  
 bladder"  
 /sex="male"  
 /tissue\_type="urinary bladder"  
 /dev\_stage="adult"  
 /lab\_host="DH10B"  
 /note="Site\_1: SalI; Site\_2: BamHI; cDNA library was  
 prepared and sequenced in Mouse Genome Encyclopedia  
 Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in  
 RIKEN. Division of Experimental Animal Research in Riken  
 contributed to prepare mouse tissues. 1st strand cDNA was  
 primed with a primer [5'  
 GAGAGAGAGAGATCAAGAGCTCTTTTCTTTTCTTTTCTTTT 3'], cDNA was  
 prepared by using trehalose thermo-activated reverse  
 transcriptase and subsequently enriched for full-length by  
 cap-trapper. cDNA went through one round of normalization  
 to Rot = 20.0 and subtraction to Rot = 370.0. Second

strand cDNA was prepared with the primer adapter of  
sequence 15' GAGAGAGATCTCTGACGTATTAATTAATCCCCCCCCCCCC  
3'. cDNA was cleaved with XhoI and BamHI. Vector: a  
modified Bluescript KS(+) after bulk excision from  
Lambda FlC I."

BASE COUNT 36 a 29 c 32 g 58 t

ORIGIN

Alignment Scores:

Pred. No.:	2.35	Length:	155
Score:	10.00	Matches:	10
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	10.31%	Indels:	0
DB:	10	Gaps:	0

US-09-854-133-586 (1-97) x BB122746 (1-155)

QY 24 LysLysGIuArgLysLysLysArgGIuArg 33  
|||  
Db 134 AAAAAAGAAAGAAAAAAGAGAGAGAGA 105

Search completed: May 11, 2003, 17:46:15  
Job time : 1486.03 secs